Planetary and Space Science

Volume 5 1961

Editor-in-Chief: MILTON GREENBERG

Regional Editors:

D. R. BATES
W. W. KELLOGG
A. I. LEBEDINSKY
D. F. MARTYN
S. K. MITRA
M. NICOLET
F. L. WHIPPLE

PERGAMON PRESS

NEW YORK · OXFORD · LONDON · PARIS

CONTENTS

NUMBER 1

M. NICOLET: Structure of the thermosphere 1

L. HARANG and J. TRÖIM: Studies of auroral echoes—I 33

H. NORINDER and E. KNUDSEN: Recent results in the investigation of the relation between lightning discharges and whistlers 46

W. E. JAYE, R. B. DYCE and R. L. LEADABRAND: Radar echoes obtained from earth satellites 1957 Alpha and 1957 Beta 50

T. OBAYASHI and Y. HAKURA: Polar ionospheric disturbances and solar corpuscular emissions 59

RESEARCH NOTES

V. A. BAILEY: Measurement of the net electric charge on the sun by means of the artificial planet "Pioneer V" 70

G. PETTENGILL and R. PRICE: Radar echoes from Venus and a new determination of the solar parallax . 71

N. I. FEDOROVA: On emission λ 10830 Å in aurorae 75

N. N. SHEFOV: On the nature of helium emission λ 10830 Å in aurorae 75

E. C. Y. INN: Origin of the *D*-region 76

SELECTED ABSTRACTS TRANSLATED FROM *ARTIFICIAL EARTH SATELLITES U.S.S.R.* 79

NUMBER 2

M. H. REES, A. E. BELON and G. J. ROMICK: The systematic behaviour of hydrogen emission in the aurora—I 87

B. HAURWITZ: Wave formations in noctilucent clouds 92

G. C. REID and M. H. REES: The systematic behaviour of hydrogen emission in the aurora—II. . . . 99

L. HARANG and J. TRÖIM: Investigation of auroral echoes—II 105

J. S. GREENHOW and J. E. HALL: The height variation of the ambipolar diffusion coefficient for meteor trails . 109

G. D. MAGNUSON and D. B. MEDVED: Sputtering as a possible mechanism for increase of ionization in the vicinity of low-altitude satellites 115

Z. SEKERA and J. V. DAVE: Determination of the vertical distribution of ozone from the measurement of diffusely reflected ultra-violet solar-radiation 122

PETER M. MILLMAN and IAN HALLIDAY: The near-infra-red spectrum of meteors 137

YALE MINTZ: Temperature and circulation of the Venus atmosphere 141

G. KVIFTE: Temperature measurement from OH bands 153

G. KVIFTE: Relative population of OH $X^2\Pi$ levels 158

RESEARCH NOTES

J. L. WEINBERG and G. NEWKIRK: Air glow of Venus: a re-examination 163

K. P. CHOPRA: Production of carbon in the upper atmosphere 164

D. E. JONES: The microwave temperature of Venus 166

NUMBER 3

J. AARONS, H. E. WHITNEY, R. S. ROGER, J. THOMSON, J. BOURNAZEL, E. VASSY, H. A. HESS, K. RAWER, N. CARRARA, P. CHECCACCI, B. LANDMARK, J. TRÖIM, B. HULTQVIST and L. LISZKA: Atmospheric phenomena noted in simultaneous observations of 1958 δ 2 (Sputnik III) 169

T. M. DONAHUE, J. E. BLAMONT and M. L. LORY: The influence of resonance absorption in the sodium layer on temperature and density measurements in the upper atmosphere 185

B. HAURWITZ: Comments on tidal winds in the high atmosphere 196
F. MÖLLER: Atmospheric water vapor measurements at 6–7 microns from a satellite 202
R. J. DAVIS and E. A. GODFREDSEN: Optimum resolving power for an ultraviolet space telescope . . 207
L. LISZA: Local electron densities deduced from the Faraday fading of satellite transmissions using measurements during two consecutive transits 213

SELECTED PAPERS AND ABSTRACTS TRANSLATED FROM *ARTIFICIAL EARTH SATELLITES U.S.S.R.*
V. I. KRASOVSKII: Soviet investigations of the ionosphere employing rockets and artificial earth satellites . 223
I. S. SHKLOVSKII: Optical methods for tracking earth satellites 233
L. I. SEDOV: Dynamic effects in the motion of artificial earth satellites 238
S. SH. DOLGINOV, L. N. ZHUZGOV and N. V. PUSHKOV: Preliminary report on geomagnetic measurements on the third Soviet artificial earth satellite 244
V. I. KRASOVSKII, YU. M. KUSHNER, G. A. BORDOVSKII, G. F. ZAKHAROV and E. M. SVETLITSKII: The observation of corpuscles by means of the third artificial earth satellite 248
ABSTRACTS 250

NUMBER 4

D. R. BATES and T. N. PATTERSON: Hydrogen atoms and ions in the thermosphere and exosphere . . 257
R. GENDRIN: Le guidage des whistlers par le champ magnetique 274
B. WARNER: The lunar maria 283
G. FIELDER: Stresses in the Moon 286
K. P. CHOPRA: Excitation of plasma waves by bodies moving in ionized atmosphere 288
K. D. COLE: Electro-hydromagnetic waves in a fully ionized gas—II 292
J. A. NILSON and G. G. SHEPHERD: Upper atmospheric temperatures from Doppler line widths—I. Some preliminary measurements on OI 5577 Å in aurora 299
W. J. MEDD and N. W. BROTEN: Lunar temperature measurements at 3200 Mc/s 307
G. V. GROVES: Methods of calculating atmospheric densities from satellite orbit data 314
B. WARNER: Accretion and erosion on the surface of the Moon 321

RESEARCH NOTES
H. NORINDER and E. KNUDSEN: The dispersion of whistlers compared with the geomagnetic latitudes of their sources 326
D. R. BATES and T. N. L. PATTERSON: An estimate of the temperature in the exosphere 328

BOOK REVIEWS

AUTHOR INDEX TO VOLUME 5

AARONS, J. 169

BAILEY, V. A. 70 (R)
BATES, D. R. 257, 328 (R)
BELON, A. E. 87
BLAMONT, J. E. 185
BORDOVSKII, G. A. 248
BOURNAZEL, J. 169
BROTEN, N. W. 307
BUTLER, H. E. 329 (B)

CARRARA, N. 169
CHECCACCI, P. 169
CHOPRA, K. P. 164 (R), 288
COLE, K. D. 292

DAVE, J. V. 122
DAVIS, R. J. 207
DOLGINOV, S. SH. 244
DONAHUE, T. M. 185
DUNGEY, J. W. 329 (B)
DYCE, R. B. 50

FEDOROVA, N. I. 75 (R)

Fielder, G. 286

Gendrin, R. 274
Godfredsen, E. A. 207
Greenhow, J. S. 109
Groves, G. V. 314

Hakura, Y. 59
Hall, J. E. 109
Halliday, I. 137
Harang, L. 33, 105
Haurwitz, B. 92, 196
Hess, H. A. 169
Hultqvist, B. 169

Inn, E. C. Y. 76 (R)

Jaye, W. E. 50
Jones, D. E. 166 (R)

Knudsen, E. 46 326 (R)
Krasovskii, V. I. 223, 248
Kushner, Yu. M. 248
Kvifte, G. 153, 158

Landmark, B. 169
Leadabrand, R. L. 50
Liszka, L. 169, 213
Lory, M. L. 185

Magnuson, G. D. 115
Medd, W. J. 307
Medved, D. B. 115
Millman, P. M. 137
Minnis, C. M. 329 (B)
Mintz, Y. 141
Möller, F. 202

Newkirk, G. 163 (R)
Nicolet, M. 1
Nilson, J. A. 299
Norinder, H. 46, 326 (R)

Obayashi, T. 59

Patterson, T. N. L. 257, 328 (R)
Pettengill, G. 71 (R)
Price, R. 71 (R)
Pushkov, N. V. 244

Rawer, K. 169
Rees, M. H. 87, 99
Reid, G. C. 99
Roger, R. S. 169
Romick, G. J. 87

Sedov, L. I. 238
Sekera, Z. 122
Shefov, N. N. 75 (R)
Shepherd, G. G. 299
Shklovskii, I. S. 233
Svetlitskii, E. M. 248

Thomson, J. 169
Tröim„ J. 33, 105, 169

Vassy, E. 169

Warner, B. 283, 321
Weinberg, J. L. 163 (R)
Whitney, H. E. 169

Zakharov, G. F. 248
Zhuzgov, L. N. 244

Planetary and Space Science

CUMULATIVE INDEXES
Volume 1–5
1959–1961

Editor-in-Chief: MILTON GREENBERG

Regional Editors:

D. R. BATES
A. I. LEBEDINSKY
S. K. MITRA
W. W. KELLOGG
D. F. MARTYN
M. NICOLET
F. L. WHIPPLE

PERGAMON PRESS

NEW YORK · OXFORD · LONDON · PARIS

Subject Index

Atmospheric structure
arctic region: **2,** 33.
deduced from satellite and rocket studies: **1,** 20, 43, 115, 136, 184, 259.
influence of resonance absorption in the sodium layer on temperature and density measurements: **5,** 185.
methods of calculating density from orbit data: **5,** 314.
relation to radiation belts: **2,** 165, 263.
temperatures from Doppler line widths: **5,** 299.
temperature measurement from hydroxyl bands: **5,** 153.
theoretical survey: **5,** 1.
Afterglows
excitation of λ 5577 in nitrogen afterglow: **3,** 169.
short duration nitrogen afterglow: **1,** 40.
spectroscopic study of nitrogen and nitrogen plus oxygen afterglows in vacuum region: **1,** 7.
Airglow, *see* nightglow and twilight airglow
Artificial airglows
morning twilight release of cesium and sodium: **2,** 174.
relative intensities of sodium lines: **2,** 71.
Artificial electron clouds: **1,** 227, 291, 306; **2,** 17, 174, 228; **3,** 139.
Atmospheric constitution
carbon: **5,** 164.
hydrogen; **5,** 257.
nitric oxide and molecular oxygen: **1,** 161.
ozone: **5,** 122.
water vapour: **5,** 202.
Atmospheric, absorption, refraction and scintillation measurements at 4700 Mc/s: **1,** 50.
Atmospheric phenomena noted in simultaneous observations of 1958 delta 2: **5,** 169.
Auroral echoes: **5,** 33, 105.
Auroral particles, magnetic separation: **2,** 49.
Auroral spectra
cause of helium emission: **5,** 75.
emission of λ 10830Å: **5,** 75.
hydrogen emission and two types of auroral spectra: **1,** 57.
narrow hydrogen emission: **2,** 130.
polarization of red oxygen line: **1,** 105; **2,** 73.
systematic behaviour of hydrogen emission: **5,** 87, 99.
temperatures from Doppler line widths: **5,** 299.
time delay of red oxygen lines: **2,** 246.
Auroral streams and solar radio bursts: **1,** 325.

Book reviews
Some ionospheric results obtained during the International Geophysical Year (W. J. G. Beynon, Ed.): **5,** 329.
The Exploration of Space (R. Jastrow, Ed.): **5,** 329.
Space Astrophysics (W. Liller, Ed.): **5,** 329.

Carbon, production in upper atmosphere: **5,** 164.
Čerenkov radiation: **1,** 105.
Contamination by extra terrestrial exploration
second meeting of the *ad hoc* committee: **1,** 249.
Continuum, bound-free of C^-: **2,** 10.
Corpuscular stream, observation by satellite: **5,** 248.
Cosmic radio noise, magnetic storms and: **2,** 99.

Cosmic rays
sidereal anistropy of high energy cosmic rays: **1,** 155; **2,** 187.
and solar radio bursts: **2,** 104.
study of cosmic ray soft component by third Soviet sputnik: **1,** 86.
COSPAR: **1,** 2.

Electro-hydromagnetic waves: **1,** 319; **5,** 292.
Energy sources in upper atmosphere: **1,** 14.
Exosphere, temperature: **5,** 328.

Gases, thermodynamic and transport properties:
intermolecular potentials for ionic systems: **3,** 217.
measurements at high temperatures: **3,** 221, 228, 238, 253, 271.
reacting gases: **3,** 238.
survey: **3,** 205.
transport properties of atomic hydrogen: **3,** 236.
Geomagnetic field
lines of force in space: **1,** 285.
measurements by satellite: **5,** 244.
origin of S_D variations: **2,** 49.

Helium-3, distribution in the Grant meteorite: **1,** 66.
Hydrogen, distribution in atmosphere: **5,** 257.
Hypersonic flight
blunt body problem, interferometric study: **4,** 337.
diffusion effects on shock structure in plasma: **4,** 358.
electrode contamination and properties of air-arc plasmas: **4,** 374.
heat transfer coefficients, measurement of: **4,** 352.

IGY, rockets and satellites: a report of the Moscow meetings August 1958: **1,** 71.
Interplanetary gas, ionization and expected emission lines: **2,** 86.
Ionization near low altitude satellites: **5,** 115.
Ionosphere
heavy absorption in the VHF-band in Arctic: **2,** 102.
local electron densities deduced from Faraday fading of satellite transmissions: **5,** 213.
radar echoes from 1957 alpha and 1957 beta: **5,** 50.
Soviet investigations: **5,** 223.
Ionosphere, E layer, geomagnetic and current
control of absorption: **2,** 1.
Ionosphere, E_s layer
characteristics: **1,** 205.
origin and geomagnetic control of: **2,** 56.
relation with spread F: **2,** 195.
Ionosphere, F layer
absorption and electron density determined from measurements of transmitted radar signals from satellites: **1,** 130.
meteoritic dust: **2,** 121.
spread F at Brisbane: **2,** 133, 150.
triple splitting at high and mid-latitudes: **2,** 214.
Ionosphere, lower
origin of D region: **5,** 76.
strongly absorbing layers below 50 km: **1,** 193.

Laboratory studies:
atomic concentration measurements by electron paramagnetic resonance: **3,** 194.
atomic oxygen, rate of recombination: **3,** 175.
chain reactions, influence of periodic pressure variations on: **3,** 94.
chemical analysis of transient species with microsecond resolution: **3,** 118.
electron attachment, recombination, elastic and inelastic scattering in atmospheric gases: **3,** 104.
electron densities in shock layers: **3,** 99.
modulated atomic beam techniques: **3,** 162.
molecules, investigations with Lozier electron impact apparatus: **3,** 187.
radiation from shock fronts in oxygen nitrogen and air: **3,** 132.
spectroscopic measurement of temperature of shock heated oxygen: **3,** 158.
thermodynamic relaxation of atmospheric gases in shock waves: **3,** 113.
vibrational relaxation in nitric oxide: **3,** 86.

Magnetic storms
cosmic radio noise and: **2,** 99.
solar radio events and: **2,** 223.
Mars, multicolor photometry: **2,** 26.
Meteor trails
ambipolar diffusion: **1,** 125; **5,** 109.
determination of winds between 80 and 100 km in medium latitudes: **1,** 94.
Meteors
fall-out of iron particles: **2,** 39.
infra-red spectrum: **5,** 137.
Meteorites, distribution of He^3 in the Grant meteorite and a determination of the original mass: **1,** 66.
Meteoritic dust: **2,** 121.
Micrometeorites,
impacts at 135 km: **2,** 75.
measurements from 1958 alpha and gamma satellities: **1,** 27.
Moon
accretion and erosion of surface: **5,** 321.
geological interpretation of surface: **2,** 256.
maria: **5,** 283.
stress history: **2,** 249.
stresses: **5,** 286.
temperature measurements at 3200 Mc/s: **5,** 307.
thermal radiation at 1420 Mc/s: **1,** 213.

Nightglow
far ultra-violet radiation: **1,** 3.
hydrogen emission: **1,** 63.
relative population of OH $X^2\Pi$ levels: **5,** 158.
temperature measurement from hydroxyl bands: **5,** 153.
Nitric oxide in upper atmosphere: **1,** 161.
Nitrous oxide, photolysis: **1,** 333.
Noctilucent clouds
nature and origin of particles: **1,** 146.
wave formations in: **5,** 92.

Organic compounds, photo-chemical formation: **2,** 79.
Oxygen, molecular, in upper atmosphere: **1,** 161.
Ozone, vertical distribution: **5,** 122.

Plasma waves, excitation by bodies moving in ionized atmosphere: **5,** 288.
Propulsion and guidance
air scooping vehicle: **4,** 159.
atomic oxygen ram-jet: **4,** 60.
clustered-rocket flights: **4,** 194.
hydrogen peroxide altitude control systems: **4,** 184.
non-chemical rockets, thermodynamic study: **4,** 111.
nuclear rockets: **4,** 145.
plasma accelerators: **4,** 102, 133.
shaped charges: **4,** 45.
spike-nozzle contour for optimum thrust: **4,** 92.
spreading of rocket exhaust jets at high altitudes: **4,** 77.
Thor, Atlas and Titan missiles: **4,** 34.

Radiation belts and properties of upper atmosphere: **2,** 165, 263.
Radiation, ionizing, detected by Pioneer II: **1,** 343.
Radio wave propagation characteristics: **2,** 110.
Rayleigh scattering by atmosphere: **1,** 265, 272.
Reactions, gas phase: *see* laboratory studies.
Reactions, surface and solid phase.
ablation of thermosetting plastics: **3,** 83.
Boltzmann equation for flows with chemical reactions: **3,** 4.
fluorocarbon resins, behaviour at high temperatures: **3,** 73.
heterogeneous pyrolysis of hydrocarbon polymers: **3,** 61.
interaction between carbon and dissociated gases: **3,** 46.
nitrogen atom recombination on surfaces: **3,** 38.
non-catalytic surface for dissociated combustion gases: **3,** 53.
oxidation of a metal forming a volatile oxide, effect of gaseous diffusion on, **3,** 31.
oxides, vaporization and thermodynamic properties: **3,** 12.
plastic laminates, behaviour at high temperatures: **3,** 68.
pyrolysis of plastics in ablation: **3,** 82.
refractory substances, rate of vaporization: **3,** 24.
Re-entry
ablating materials: **4,** 447, 463, 479.
dynamics and thermodynamics: **4,** 382.
epicyclic techniques for flight test data analysis: **4,** 486.
heat conduction in a slab with sublimation at surface: **4,** 436.
lifting vehicles: **4,** 399.
manned: **4,** 411, 422.
Sputnik I rocket: **1,** 43.
Rockets, *see* vehicles.
Rocket motors:
combustion chambers, fabrication techniques for solid propellant: **4,** 14.
non-destructive testing: **4,** 1.

Satellites, *see* vehicles.
Satellites
auxiliary power systems, survey of requirements: **4,** 202.
decay of spin: **1,** 43.
dynamic effects in motion of: **5,** 238.
ionization near low altitude: **5,** 115.
lifetimes: **2,** 212.
optical methods of tracking: **5,** 233.
polonium-fuelled thermo electric generator (SNAP III): **4,** 226.
radar echoes from 1957 alpha and 1957 beta: **5,** 50.

S_D field, origin: **2,** 49.
Solar parallax, new determination: **5,** 71.
Solar radio bursts
association with auroral streams: **1,** 325.
cosmic rays: **2,** 104.
Sun, net electric charge on: **5,** 70.

Telescope, space: **5,** 207.
Thermosphere structure: **5,** 1.
Twilight airglow
influence of resonance absorption in the sodium layer on temperature and density measurements: **5,** 185.
source of N_2^+ ions: **2,** 157.

URSI, report on the symposium on the use of space vehicles at the fall 1958 meeting: **1,** 238.

Vehicles
beam vibration testing: **4,** 285.
circular pitching and yawing motion of nose cone configurations: **4,** 328.
comparison of structural responses of a captive missile with a free-free missile: **4,** 301.
dynamic loads from fuel motion in missile tanks: **4,** 313.
effect of cost on design and size of multi-stage rockets: **4,** 242.
low cost satellite launched from aircraft: **4,** 249.
spin stabilization system for Able space probe: **4,** 262.
see also hypersonic flight, propulsion and guidance re-entry, rocket motors, satellites, auxiliary power systems.
Venus
airglow: **1,** 32; **5,** 163.
micro-wave temperature: **5,** 166.
radio echoes from: **5,** 71.
temperature and circulation in atmosphere: **5,** 141.
VLF emission
associated with magnetic disturbance: **1,** 253.
observations at Australian stations: **2,** 238.

Water vapor in atmosphere, measurements at 6-7 microns from a satellite: **5,** 202.
Waves
electro-hydromagnetic: **1,** 319, **5,** 292.
plasma, excitation by bodies moving in ionized atmosphere: **5,** 288.
Whistlers
current jet hypothesis: **2,** 72.
dispersion, relation with geomagnetic latitude of source; **5,** 326.
guidance by geomagnetic field: **5,** 274.
penetration through ionospheric layers: **2,** 261.
relation with lightning discharges: **1,** 173; **5,** 46.
Winds
between 80 and 100 km in medium latitudes (meteor train radio study): **1,** 94.
tidal: **5,** 196.

Author Index

(B signifies book review and R research note)

Aarons, J. **1**, 50; **5**, 169
Abrams, I. J. **4**, 34
Abramson, H. N. **4**, 313
Ackermann, R. J. **3**, 12
Ainsworth, J. **2**, 33
Allport, J. J. **3**, 86.
Ambrosio, A. **4**, 436
Amdur, I. **3**, 228
Anderson, B. H. **4**, 352
Anderson, T. P. **4**, 77
Aschenbrand, L. M. **1**, 227, 291, 306; **2**, 17, 174, 228; **3**, 139

Baer, D. **4**, 436
Bailey, V. A. **5**, 70 (R)
Bates, D. R. **5**, 257, 328 (R)
Barrett, W. F. **4**, 411
Barron, S. **4**, 45
Barth, C. A. **3**, 194
Bauer, S. H. **3**, 118
Belon, A. E. **5**, 87
Benson, S. W. **3**, 257
Berner, F. **4**, 159
Bhonsle, R. V. **2**, 99
Biondi, M. A. **3**, 104
Blamont, J. E. **5**, 185
Booth, F. B. **3**, 194
Bordovskii, G. A. **5**, 248
Bortner, M. H. **3**, 99
Bossolasco, M. **1**, 205
Bournazel, J. **5**, 169
Bowman, G. G. **2**, 133, 150, 195, 214
Brackman, R. T. **3**, 162
Breene, R. G. **2**, 10
Brice, N. M. **2**, 238
Brokaw, R. S. **3**, 238
Broten, N. W. **5**, 307
Brown, R. **4**, 313
Bryan, C. R. **4**, 486
Buckingham, R. A. **3**, 205
Burgers, J. M. **3**, 4
Buss, J. H. **3**, 257
Butler, H. E. **5**, 329 (B)
Byram, E. T. **1**, 3

Camac, M. **4**, 159
Carrara, N. **5**, 169
Casey, J. **1**, 50
Castelli, J. P. **1**, 50
Chamberlain, J. W. **2**, 60, 73 (R), 157
Champion, K. W. **1**, 259
Checcacci, P. **5**, 169
Chopra, K. P. **5**, 164 (R), 288
Chubb, T. A. **1**, 3
Chudakov, A. E. **1**, 86
Clark, K. C. **3**, 169
Cleary, G. F. **4**, 249
Cohen, J. B. **3**, 68
Cole, K. D. **1**, 319; **5**, 292
Coleman, Jnr., P. J. **1**, 343
Coulson, K. L. **1**, 265, 277
Crouchley, J. **2**, 238
Curtiss, C. F. **3**, 204
Cutting, J. C. **3**, 53

Dalgarno, A. **3**, 217, 236
Dattilo, A. W. **4**, 45
Dave, J. V. **5**, 122
Davis, R. J. **5**, 207
Deirmendjian, Diran **1**, 146
Diaconis, N. S. **4**, 463
Dolan, L. P. **4**, 1
Dolginov, S. Sh. **5**, 244
Donahue, T. M. **5**, 185
Dubin, M. **2**, 121
Duncan, R. A. **1**, 112
Dungey, J. W. **5**, 329 (B)
Dyce, R. B. **5**, 50

Efimenko, J. **3**, 24
Elena, A. **1**, 205
Elford, W. G. **1**, 94
Ellis, G. R. A. **1**, 253
Epperson, T. B. **4**, 313
Escobar, I. **1**, 155: **2**, 187

Fanucci, J. B. **4**, 463
Favin, S. **3**, 221
Fay, J. A. **3**, 53
Fedorova, N. I. **5**, 75 (R)
de Feiter, L. D. **2**, 223
Ferioli, C. **1**, 50
Fielder, G. **5**, 286
Fineman, M. A. **3**, 187
Fireman, E. L. **1**, 66
Fite, W. L. **3**, 162
Fokker, A. D. **2**, 223
Friedman, H. **1**, 3
Friel, P. J. **3**, 83

Galman, B. A. **4**, 399
Galperin, G. I. **1**, 57
Garmire, G. **1**, 325
Gendrin, R. **5**, 274
Gloersen, P. **4**, 102
Godfredsen, E. A. **5**, 207
Goldbaum, G. C. **4**, 242
Golden, J. A. **3**, 175
Gorchakov, E. V. **1**, 86
Gorowitz, B. **4**, 102
Greenberg, O. W. **4**, 358
Greenhow, J. S. **5**, 109
Gross, L. W. **4**, 226
Groth, W. E. **1**, 333; **2**, 79
Groves, G. V. **5**, 314

Hakura, Y. **5**, 59
Hall, J. E. **5**, 109
Halliday, I. **5**, 137
Hammerling, P. **3**, 132
Harang, L. **1**, 102; **5**, 33, 105
Harris, I. **1**, 20, 37
Haurwitz, B. **5**, 92, 196
Hess, H. A. **5**, 169
Hildebrant, A. F. **3**, 194
Hogan, W. T. **3**, 53
Hoffmann, W. C. **1**, 238; **2**, 72 (R)
Holister, G. S. **3**, 162
Horowitz, R. **2**, 33
Huffman, F. N. **4**, 226
Hultqvist, B. **1**, 193; **5**, 169
Hurwicz, H. **4**, 447

Inn, E. C. Y. **1**, 7; **5**, 76 (R)

Jastrow, R. **1**, 20, 37
Jaye, W. E. **5**, 50
Jelley, J. V. **1**, 105
Jones, D. E. **5**, 166 (R)
Jordan, P. F. **4**, 411
Jursa, A. S. **1**, 7, 161

Kahl, G. D. **4**, 337
Kazantsev, A. N. **1**, 130
Kellogg, W. W. **1**, 71
Kiefer, J. **3**, 118
Kissell, K. E. **4**, 111
Kivel, B. **3**, 132
Knudsen, E. **1**, 173; **5**, 46, 326 (R)
Krasovskii, V. I. **1**, 14; **5**, 223, 248
Kreiken, E. A. **2**, 39
Kretschmer, C. B. **4**, 60
Kuebler, N. A. **3**, 61
Kupperian, Jr., J. E. **1**, 3
Kushner, Yu. M. **5**, 248
Kvifte, G. **5**, 153, 158

Lagow, H. E. **2**, 33
Landmark, B. **5**, 169
Latvala, E. K. **4**, 77
Leadabrand, R. L. **5**, 50
LeBlanc, F. J. **1**, 7, 161
Levoy, M. M. **4**, 145
Liebman, M.E. **4**, 1
Lindberg, H. E. **4**, 285
Linder, B. **4**, 479
Linevsky, M. J. **4**, 374
Liszka, L. **5**, 169, 213
Lofgren, N. L. **3**, 24
Logachev, J. L. **1**, 86
Loh, W. H. T. **4**, 382
van Lohuizen, H. P. Th. **2**, 223
Lory, M. L. **5**, 185
Lovering, J. F. **2**, 75 (R)
Lundbak, A. **2**, 212

MacDonald, G. J. F. **2**, 249
McGinn, J. H. **4**, 374
Magnuson, G. D. **5**, 115
Malville, J. M. **2**, 130
Manring, E. R. **1**, 27; **2**, 174
Margrave, J. L. **3**, 3
Marmo, F. F. **1**, 227, 291, 306; **2**, 17, 174, 228; **3**, 139
Marsaglia, G. **4**, 194
Martinek, F. **3**, 271
Maxwell, A. **1**, 325; **2**, 104
Medd, W. J. **5**, 307
Medved, D. B. **5**, 115
Mezger, P. G. **1**, 213
Millman, P. M. **5**, 137
Minnis, C. M. **5**, 329 (B)
Mintz, Yale **5**, 141
Mintzner, R. A. **1**, 259
Miyamoto, S. **2**, 256
Modisette, J. L. **3**, 31
Moffatt, W. C. **3**, 53
Möller, F. **5**, 202
Monchick, L. **3**, 221
Monson, P. R. **3**, 86
Moses, K. **4**, 102
Murphy, Jr., C. H. **4**, 328
Murray, E. L. **1**, 125
Myers, H. **3**, 257
Myerson, A. L. **3**, 175

Nerurkar, V. N. **1**, 155; **2**, 187
Nelson, L. S. **3**, 61
Newgard, J. J. **4**, 145
Newkirk, G. **1**, 32; **5**, 163 (R)
Nicolet, M. **5**, 1
Nilson, J. A. **5**, 299
Norinder, H. **1**, 173; **2**, 261 (R); **5**, 46, 326 (R)

Obayashi, T. **5**, 59
Oldenberg, O. **1**, 40
Omholt, A. **2**, 246
Ortner, J. **1**, 193

Paetzold, H. K. **1**, 115
Palmquist, Jr. N. B. **4**, 202
Patterson, T. N. L. **5**, 257, 328 (R)
Penner, S. S. **3**, 85
Petrocelli, A. **3**, 187
Pettengill, G. **5**, 71 (R)
Pressman, J. **1**, 227, 291, 306; **2**, 17, 174, 228; **3**, 139
Price, R. **5**, 71 (R)
Prok, G. M. **3**, 38
Pushkov, N. V. **5**, 244

Ramanathan, K. R. **2**, 99
Rao, G. V. R. **4**, 92
Rawer, K. **5**, 169
Rees, M. H. **5**, 87, 99
Reid, G. C. **5**, 99
Reiter, G. S. **4**, 262
Robben, F. **3**, 86
Roberts, L. **4**, 422
Roger, R. S. **5**, 169
Rol, N. **3**, 118
Romick, G. J. **5**, 87
Rosen, A. **1**, 343
Roosen, J. **2**, 223
Ruby, S. **3**, 68
Rundle, H. N. **2**, 86

Sagan, C. **2**, 157
Sanscrainte, W. **4**, 184
Scheur, E. M. **4**, 34
Schierholz, H. **1**, 333
Schilling, G. F. **1**, 136
Schock, A. **4**, 133
Schryer, D. R. **3**, 31
Sedney, R. **4**, 337
Sedov, L. I. **5**, 238
Sekera, Z. **5**, 122
Sen, H. K. **4**, 358
Settlage, P. H. **3**, 73
Shaw, J. E. **2**, 1, 49, 56
Sheehan, W. F. **4**, 262
Shefov, N. N. **5**, 75 (R)
Shepherd, G. G. **5**, 299
Sherman, A. **3**, 271
Shklovskii, I. S. **1**, 63; 5, 233
Sibley, W. L. **1**, 285
Siegle, J. C. **3**, 73
Simha, R. **3**, 82
Singer, S. F. **2**, 165
Siry, J. W. **1**, 184
Skinner, G. T. **3**, 253
Sonett, C. P. **1**, 343
Spindler, R. J. **4**, 447
Steverding, B. **3**, 113
Strassl, H. **1**, 213
Sutton, G. W. **4**, 463
Svetlitskii, E. M. **5**, 248

Tanaka, Y. **1**, 7, 161
Teare, J. D. **3**, 132
Thompson, A. R. **1**, 325; **2**, 104
Thomson, J. **5**, 169
Thorn, R. J. **3**, 12
Treanor, C. E. **3**, 253
Treve, Y. M. **4**, 358
Tröim, J. **1**, 102; **5**, 33, 105, 169

Vakulov, P. V. **1**, 86
Vassy, A. **2**, 71 (R)
Vassy, E. **2**, 71 (R); **5**, 169
de Vaucouleurs, G. **2**, 26
Vernov, S. N. **1**, 86
Vestine, E. H. **1**, 146, 285

Wachmann, H. Y. **4**, 374
Walker, R. E. **3**, 221
Walker, R. F. **3**, 24
Wallace, L. **2**, 60
Ward, F. K. **3**, 68
Warner, B. **5**, 283, 321
Warwick, J. W. **1**, 43
Weil, R. **1**, 155; **2**, 187
Weinberg, J. L. **5**, 163 (R)
Werner, L. **3**, 113
Westenberg, A. A. **3**, 221
Weyssenhoff, H. v. **2**, 79
White, J. F. **4**, 242
Whitney, C. A. **1**, 136
Whitney, H. E. **5**, 169
Wight, H. M. **3**, 94
Winslow, G. H. **3**, 12
Wood, J. D. **4**, 301
Wurster, W. H. **3**, 158

Young, R. A. **3**, 169

Zakharov, G. F. **5**, 248
Zelikoff, M. **3**, 138
Zhekulin, L. **2**, 110
Zhuzgov, L. N. **5**, 244
Zinman, W. **3**, 46
Zukerman, A. **4**, 60
Zwissler, L. E. **4**, 14

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 1-32. Printed in Great Britain.

STRUCTURE OF THE THERMOSPHERE

MARCEL NICOLET

Ionosphere Research Laboratory, Pennsylvania State University, University Park, Pennsylvania

(*Received* 3 *July* 1960)

Abstract—The vertical distribution of the density in the thermosphere, deduced from satellite observations, must be explained by an increase of the scale height with altitude. A varying gradient of the scale height cannot be interpreted by assuming an increase of the temperature gradient with altitude. An examination of the interrelationships between the absolute values of density in a dark atmosphere and diurnal conditions of heat conduction reveals that the varying gradient of the scale height above 200 km is essentially due to the decrease of the molecular weight, mg, of the atmospheric constituents subject to diffusion.

In the night atmosphere the isothermy above a certain altitude ($>$200 km) is the critical factor characterizing the vertical distribution of density. The temperature of the isothermal region, resulting from conduction, is related to the ultra-violet heating which was available during the day. The effect of diffusion has been clearly shown by establishing a thermo-isobaric relation connecting the temperature of the isothermal region with an isobaric level where atomic oxygen has a specific concentration. From observational data on the variation of the night-time density at high levels, it is possible to deduce the variation of the temperature of the isothermal region.

The gradient of temperature in a sunlit atmosphere is related to the fraction of the ultra-violet Solar energy absorbed, which determines the magnitude of the variation of the scale height with altitude. Since heat transport is a function of the atomic or molecular concentrations and the square of the distance, it is shown that anomalies in the temperature gradient cannot be permanent.

1. INTRODUCTION

The effect of air resistance on the motion of an artificial earth satellite makes it possible to derive the atmospheric density in the region of the perigee of its orbit. Formulae relating satellite drag to orbital elements have been derived by various authors; see for example Groves[20], King-Hele[38] and Sterne[80]. Determinations of density have been made by various authors following the initial calculations of the acceleration of the first two satellites 1957 α and β (Sputniks 1 and 2), whose perigees were at about 220 km; see for example Mullard Radio Astronomy Observatory[53], Royal Aircraft Establishment[70], Sterne *et al.*[82], Sterne and Schilling[81], Harris and Jastrow[22], Jacchia[26,29], Groves[17], Sterne[78,79], El'Yasberg[16], Lidov[47], Mikhnevich[52], Priester *et al.*[68], Warwick[84] and Paetzold[63].

From the spring of 1958, the satellites Vanguard I (1958 β2), Explorer I (1958 α), Explorer IV (1958 ε), and Sputnik III (1958 δ), made possible an analysis of the densities at altitudes of approximately 650 km, 350 km, 260 km and 220 km corresponding to the perigees of the above satellites. See for example, Jacchia[27], Jacchia and Briggs[34], Harris and Jastrow[23], Siry[77], Sterne[78,79], Schilling *et al.*[74], Schilling and Whitney[72,73], Sedov[75], King-Hele[37], Groves[19,20,21], Mikhnevich *et al.*[52] and Paetzold[65].

It is evident that the *absolute* values of the density which have been deduced from these observations may vary according to the methods and satellite parameters used by the various authors. Thus, the drag parameter η_s is dependent on the mass of the satellite m_s, on the effective cross-section of the satellite s_s, and on the satellite's drag coefficient C_s; $\eta_s = s_s C_s / m_s$ can differ according to the values used by the authors. Consequently, the absolute values of the density deduced can be different even if the formulae relating density to acceleration are the same. In addition, the values of the density do not necessarily relate to the same period in the life of a satellite and can correspond, therefore, to different physical states of the high atmosphere. In this case, the scale height associated with the density may not be appropriate. Finally, because the altitudes of certain satellites

are not given with adequate precision the densities obtained do not correspond with the approximate altitude indicated.

From the beginning of these observations, Groves[17] stressed the importance of the equatorial bulge which changes the altitudes by 12·5 km between the equator and latitude 50°. After Jacchia[27] had indicated the existence of irregularities in the acceleration of satellite 1957 β many variations were detected and ascribed to many different causes but it was immediately obvious that the atmosphere was responsible for these irregularities. Nevertheless, the special effects were attributed to many causes such as discontinuities in the atmosphere at certain latitudes. For example, King-Hele and Walker[40] insisted that near 30°N the effect of the irregularities was caused mainly by solar disturbances. In any event, the variations of solar radiations at 20 cm and 10 cm (Priester[67], Jacchia[30]) show clearly that solar emissions play a primary role in the variation of atmospheric density. It should be noted here that it is only possible to enter into a detailed discussion of all the variations if the observational data are very precise and sufficient in number so as to be able to follow all the fluctuations as a function of altitude.

2. ANALYSIS OF OBSERVATIONAL RESULTS

2.1. *Mean values of the density*

Before carrying out an analysis of the various variations of the density it is desirable to provide a description of the results as a whole. For this reason we have summarized the main results as shown in Fig. 1.

Before being able to determine which variations modify the density we must decide on the average conditions. The different determinations shown in Fig. 1 were made for different periods, although most of them relate to the beginning of 1958. It should be noted that the rocket results, for altitudes of the order of 200 km, give variations which do not appear to agree with the results deduced from the acceleration of satellites. It appears, however, that the rocket measurements of density carried out at about 200 km lead to values of $(4 \pm 2) \times 10^{-13}$ g cm^{-3}, although certain differences can be expected since measurements made with rockets involve the collection of samples which can relate to sporadic conditions that are not necessarily representative. As a result, values of density between 200 and 250 km as given by Mikhnevich[51] are perhaps too small when compared with satellite data.

If we consider the results of LaGow *et al.*[46] and Horowitz *et al.*[25], the densities, ρ, at 200 km, at Churchill, i.e.

1956, Nov. 17, day; $\rho = 3{\cdot}6\,(^{+3}_{-1{\cdot}5}) \times 10^{-13}$ g cm^{-3}
1957, July 29, day; $\rho = (6{\cdot}7 \pm 2) \times 10^{-13}$ g cm^{-3}
1958, Oct. 31, day; $\rho = 4{\cdot}0 \times 10^{-13}$ g cm^{-3} (extrapolated)
1958, Feb. 24, night; $\rho = (1{\cdot}3 \pm 0{\cdot}6) \times 10^{-13}$ g cm^{-3}

it is clear that the median value corresponds to the value obtained from the satellite data, $(4 \pm 2) \times 10^{-13}$ g cm^{-3}. Furthermore, the value obtained at White Sands, of the order of $1{\cdot}4 \times 10^{-13}$ g cm^{-3} at 200 km, corresponds to an atmosphere, in 1951, for which the pressure was only 10^{-4} mm Hg at 100 km when it was $(3 \pm 1) \times 10^{-4}$ mm Hg at Churchill.

Mikhnevich[51] indicating a density of $2{\cdot}7 \times 10^{-13}$ g cm^{-3} at 200 km is still inside of the possible variation. However, he adopts[52] the value of $2{\cdot}1 \times 10^{-13}$ g cm^{-3} at 225 km for an atmospheric model.

It can be concluded that an average value of the order of 4×10^{-13} g cm^{-3} represents the atmospheric density at 200 km during the sunspot maximum in 1958–1959. Variations leading to $(4 \pm 2) \times 10^{-13}$ g cm^{-3} can be accepted. Extreme values between 1 and 7×10^{-13} g cm^{-3} are not representative of latitudinal, seasonal or diurnal variations at 200 km, but should be associated with the effect of solar activity if they are not included in the possible errors of measurement.

It must be pointed out that an analysis of the varying conditions at 200 km is to be related to the analysis of conditions between 100 km and 150 km. For example, there is a variation of the density of molecular oxygen by a factor of 3 according to the measurements made by Byram *et al.*[6] and Kupperian *et al.*[45]. Like-

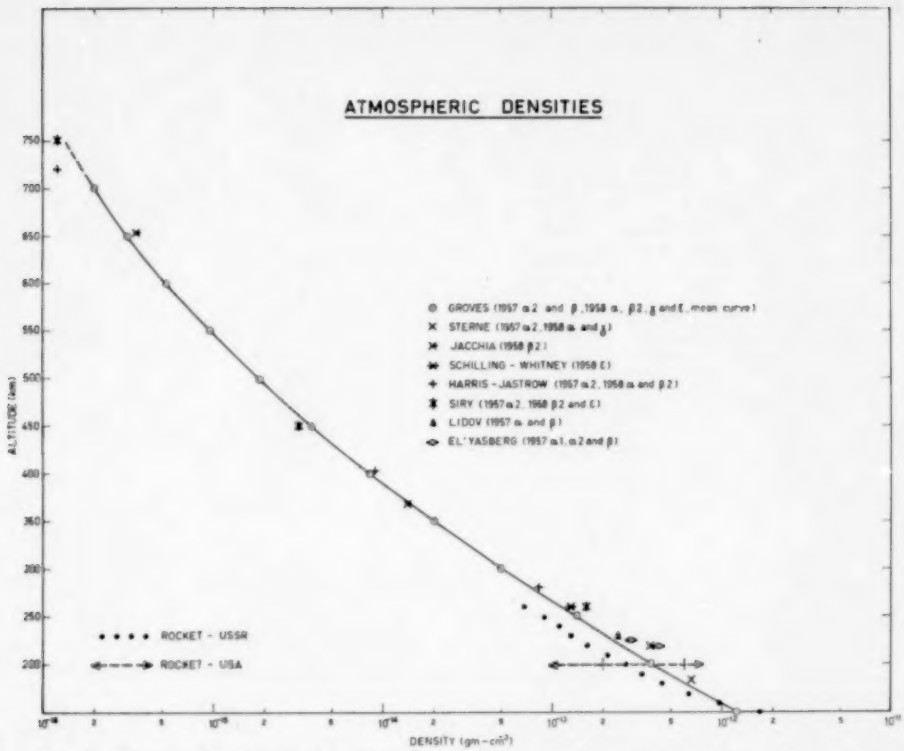

ATMOSPHERIC DENSITIES
GROVES (1957 α2 and β, 1958 α, β2, γ and ε, mean curve)
STERNE (1957 α2, 1958 α and γ)
JACCHIA (1958 β2)
SCHILLING - WHITNEY (1958 ε)
HARRIS - JASTROW (1957 α2, 1958 α and β2)
SIRY (1957 α2, 1958 β2 and ε)
LIDOV (1957 α and β)
EL'YASBERG (1957 α1, α2 and β)
ROCKET - USSR
ROCKET - USA
ALTITUDE (km)
DENSITY (gm-cm⁻³)

Fig. 1. Density–altitude relation above 150 km from various determinations. The average curve represented by Groves's data shows that the height variation of the density can only be explained by an increase of the scale height with altitude. Measurements made in 1957 and first months of 1958.

wise considering the results obtained by Horowitz and LaGow[24] at White Sands and by Horowitz *et al.*[25] at Ft. Churchill it is clear that there is a broad range of a factor 4 in the pressure and density data at 100 km. In other words if the value $(2{\cdot}5 \pm 1{\cdot}5) \times 10^{-4}$ mm Hg is accepted for the pressure at 100 km, the possibility exists that the density at 200 km is subject to a variation of 50 per cent even if the structure of the atmosphere above 100 km is essentially the same. Therefore, the data available on atmospheric density obtained by means of rockets and satellites show that variations of the density near 200 km must be connected with the atmospheric structure in the entire region between 100 and 200 km. A certain variation of the density at 200 km must be explained by the boundary conditions near 100 km and by the atmospheric structure above the later altitude.

The curve in Fig. 1 has been drawn following the determination of Groves[19] and can be considered as an average distribution for the first six months of 1958 corresponding to a certain sunlit atmosphere. The absolute values near 200 km depend on the value which is assumed for the scale height at that altitude. For example, the results of Lidov[47] lead to values for ρ between $2{\cdot}4 \times 10^{-13}$ g cm^{-3} to $3{\cdot}2 \times 10^{-13}$ g cm^{-3} if the scale heights are 50 km and 30 km, respectively. El'Yasberg[16] has adopted $H = 25$ km at 225 km. The average value of Groves[19] at 200 km is 46 km. Such differences show that variations do occur in the atmosphere above 200 km. But the main conclusion to be drawn from the average values of

the density is that the vertical distribution must be explained by an increase of the scale height H with altitude. In fact the equation for a perfect gas and the static equation indicate[57] that the density can be expressed by the relation

$$\frac{d\rho g}{\rho g} = -\frac{1+\beta}{\beta}\frac{dH}{H} \qquad (2.1)$$

where $\beta = dH/dz$ is the gradient of the scale height and g is the gravitational acceleration. The integration of equation (2.1) for a height interval sufficiently small so that β can be considered practically constant leads to

$$\frac{\rho g}{\rho_0 g_0} = \exp\left[-\frac{(1+\beta)z}{\frac{1}{2}(H+H_0)}\left\{1+\tfrac{1}{3}\left(\frac{\beta z}{H_0+H}\right)^2+\ldots\right\}\right] \qquad (2.2)$$

In using equation (2.2) we can see that H increases regularly with altitude between 150 km and 700 km. Thus the results of Groves[39], corresponding to a certain sunlit atmosphere in 1958, coincide with an average model presented by Nicolet[58]. Subsequent calculations carried out by Nicolet[59, 60] involving an analysis of the physical conditions of an atmosphere in which the temperature is constant above 220 km lead to the important conclusion that: *the vertical distribution of the density of the high atmosphere can be explained if the atmosphere in which the various constituents are subject to diffusion is isothermal above a certain level.* Such an atmospheric model shows that ultraviolet radiation is involved in the heating of the atmosphere.

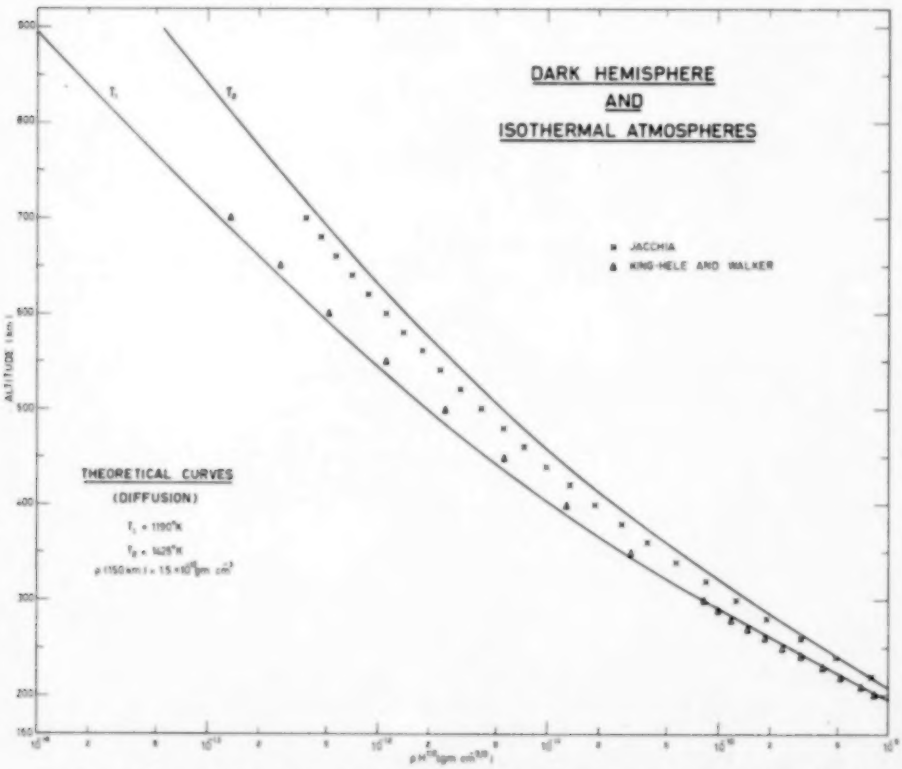


Fig. 2. $\rho H^{1/2}$ proportional to orbital acceleration of satellites in an isothermal atmosphere subject to diffusion. The density–altitude relations deduced from observational data can be followed. It is possible to fit the satellite observations of density, corresponding to a night-time atmosphere by an isothermal atmosphere.

A recent analysis made by King-Hele and Walker[42] leads to a night-time distribution in 1959 shown in Fig. 2. This figure further shows that an isothermal atmosphere at about 1200°K with diffusion beginning at 150 km can represent the observations. The observational results do not differ from the model computed by Nicolet[60]. In the same way, an analysis of night-time conditions by Jacchia[33] (see Fig. 2) also shows the possibility of following the atmospheric distribution for high solar activity when a temperature of the order of 1400°K is adopted for the isothermal layer.

The above results lead to the conclusion that there is a thermopause and that its level is subject to a diurnal variation. Its altitude is maximum in a sunlit atmosphere and is minimum in a dark atmosphere. However, when the scale heights corresponding to isothermal atmospheres in diffusion equilibrium are compared (Fig. 3) with the empirical scale heights deduced by King-Hele and Walker[42] and by Jacchia[33] considerable differences are found. King-Hele and Walker have introduced a negative gradient below 250 km and Jacchia's curve corresponds to an important gradient even at 600 km. Such differences show that arbitrary atmospheric models can be made to follow

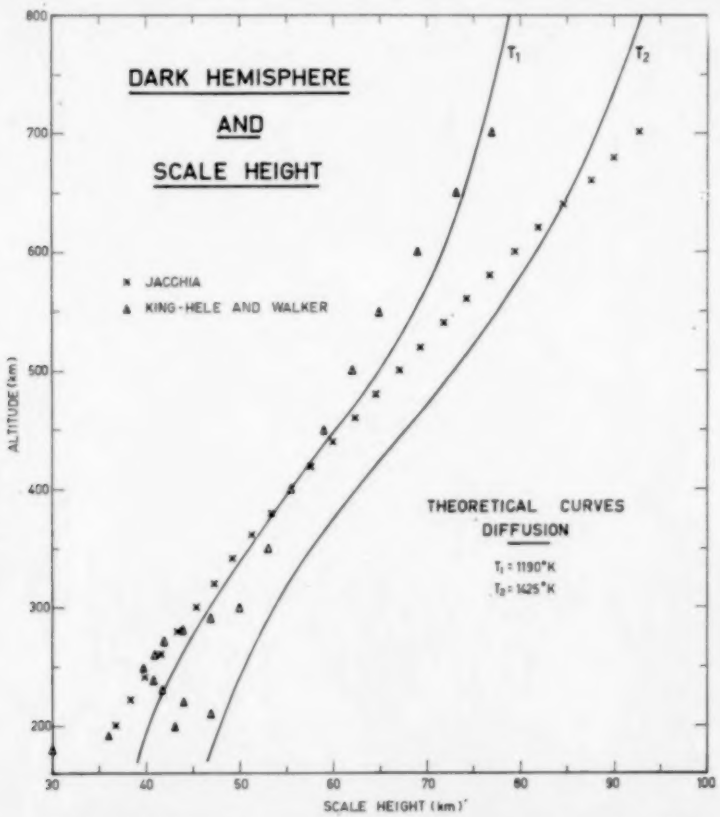


Fig. 3. Variations of the scale height with altitude. The variation of the scale height is associated with a variation of the mean molecular mass depending on diffusion. In this figure the scale height–altitude relations must be associated with the density distributions shown in Fig. 2. It is possible to deduce from observational data of $\rho H^{1/2}$ inconsistent values of the scale height.

atmospheric densities deduced from satellite observations.

2.2. *Variations of the density*

Jacchia[32] has shown that the amplitude of the fluctuations of the accelerations, proportional to $\rho H^{1/2}$, increases with the altitude of the perigee of the satellite. In fact, the important fluctuations appear simultaneously at all altitudes and the diurnal variations are magnified between 350 and 650 km.

The effect of the earth's equatorial bulge is evident. Each transit at the equator of Explorer I corresponds to a maximum and the change in perigee height of Sputnik III leads to a variation in the acceleration.

In Fig. 4 we consider data from satellite

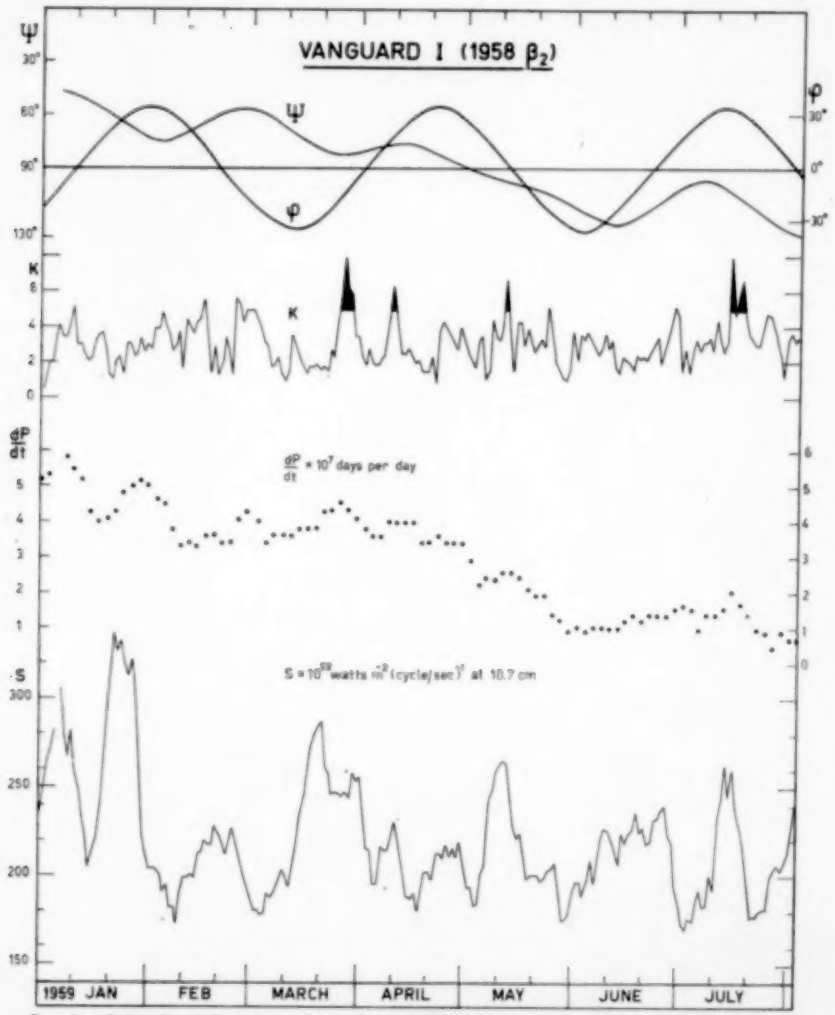


Fig. 4. Orbital acceleration of Vanguard I (1958 β2). $-dP/dt \propto \rho H^{1/2}$[5]. Perigee latitudes ($\varphi$) and angular distance of the perigee from the subsolar point (Ψ)[32]. The solar activity is represented by daily values of solar radiation at 10·7 cm (Ottawa, Covington) and the magnetic activity by the daily *K* values deduced from the three-hourly indices[2].

1958 β 2 during the first half of 1959, obtained from the values determined by Jacchia[32] and Briggs[5]. The solar activity is represented by daily values of solar radiation at 10·7 cm as observed in Canada (National Research Council) and by the daily values deduced from the three-hourly K indices provided by Bartels. It is obvious, as Jacchia has shown[32], that there is a very close correlation between the variations of the density and those of the solar radiation as obtained using the electromagnetic radiation at 10·7 cm. As for a magnetic activity relation to the corpuscular radiation, there is no general correspondence although, in certain cases, it is possible to see that corpuscular effects can occur during important geomagnetic disturbances. It appears, however, that a greater resolution in the orbital acceleration is needed to detect short-lived perturbations.

On April 30 the density has decreased (Fig. 4) when the sun is at 90° (the angle at the centre of the earth determined by the position of the sun and of the perigee of the satellite); that is, when the earth below the perigee is no longer sunlit. Finally, on July 31 the nocturnal effect is complete and the density is very low. These data show that the atmosphere at an altitude of 650 km is subject to a very important diurnal variation and the different analyses of Jacchia[32], Wyatt[56] and of Priester and Martin[69] show that the diurnal effect is the principal one.

This strong diurnal variation is confirmed by the observed accelerations of Vanguard II 1959 α[33], for which the diurnal variation of $\rho H^{1/2}$ is of the order of a factor of 5 between March and August 1959 (Fig. 5). It is therefore necessary to consider that the altitude of the thermopause and *the temperature of the isothermal atmosphere vary considerably between night and day*. Under such conditions the *normal heating of the upper atmosphere* takes place by electromagnetic radiation. It is, therefore, not necessary to look for a corpuscular effect[2] when investigating such a normal heating effect of the atmosphere, a hydromagnetic effect[14, 15] or any effect other than that of electromagnetic radiation, because such effects cannot be associated with a diurnal variation whose character is so pronounced at 650 km. Such effects must be linked with disturbances.

The variations of $\rho H^{1/2}$ at an altitude of about 350 km, corresponding to the perigee of Explorer I (1958 α), are considerably smaller than at an altitude of 650 km. From the beginning of January 1959, when the angular distance of the sun reaches 180° until it reaches 90° at the March equinox, the density decreases slightly in agreement with the diminution in solar flux. Between March and August 1959 the increase in the acceleration occurs with the fluctuations clearly associated with the sequence of perigees at the equator even if some association can be made with the solar activity.

Thus, at an altitude of 350 km, the effect of solar heating is evident and leads to a clear diurnal variation. The effect of the earth's equatorial bulge is of the order of 10 per cent.

When an analysis of the variations of the density is carried out at altitudes less than 300 km it is found that the diurnal effects are greatly diminished. In the analysis of the density deduced from the acceleration of 1958 ε[73] the diurnal effects that can be inferred for the period between July and November 1958 reach a maximum of 20 per cent; at the same time, an apparent latitude effect, caused by a change in perigee height due to the earth's equatorial bulge, is certainly of the order of 20 per cent. The discontinuities in the density at 50°N and S cannot be attributed to a latitudinal variation in air density.

The analysis of the fluctuations of the Sputniks by various authors already mentioned, in particular King-Hele[38], have shown that the diurnal effect is generally masked by other effects. In Fig. 6 the data given by Kozai[43] for 1958 δ 2 are shown. It may be seen that from January to November 1959 the maximum variation is $\pm$60 per cent. The main character of this variation corresponds to a displacement of the perigee from 24°S on 1 January 1959 to 65°S, at the beginning of June, and to less than 10°S at the end of November 1959. Thus, the perigee was located in the Antarctic during the winter. Because of this, the altitude of the perigee did not diminish with time[49]. In June its altitude, about 225 km, was a maximum and it is necessary to consider an increase in altitude

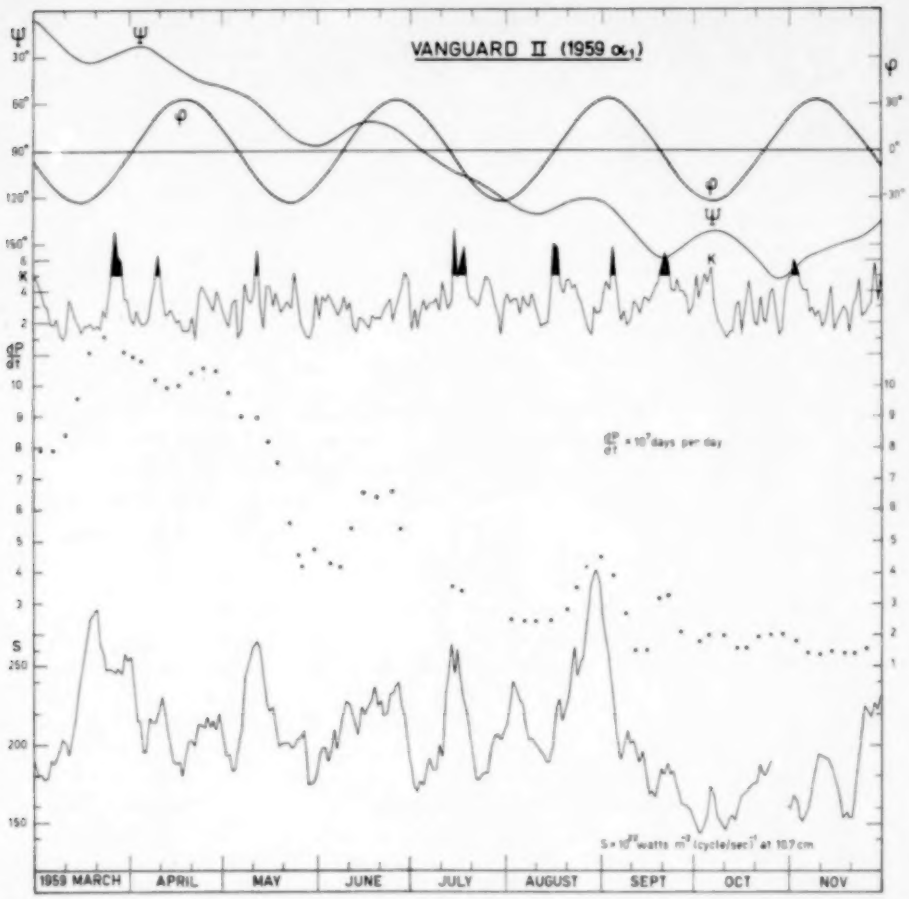


Fig. 5. Orbital acceleration of Vanguard II (1959 α). $-dP/dt \propto \rho H^{1/2}$ [33]. Other symbols, see Fig. 4.

of some 18 km with regards to the equator. As a consequence a decrease (30 per cent to 40 per cent) of dP/dt (see Fig. 6) can be explained by the flattening of the earth. It is, therefore, clear that the general variation observed in the acceleration of this satellite is due to change in the altitude of the perigee and not to a discontinuity caused by a latitude effect.

Another remarkable result is the very close association with geomagnetic storms that Jacchia[31] discovered for satellite 1958 δ 1, thanks to a resolution of 10 revolutions in the analysis of the acceleration. Fig. 7 shows the variation of the acceleration of 1958 δ 1 as given by Jacchia[31] with two remarkable peaks on 9–10 July and 4–5 September when a resolution of 10 revolutions is used in the analysis of observations. It is clear that such an increase of the atmospheric density is associated with the magnetic activity defined by the K indices. A remarkable fact that Jacchia[31] has found is that the latitudes of the perigee of 1958 δ 1 were 35°N and 15°N in 9 July and 4 September, respectively. Furthermore, it is of interest to note that the angular distances, Ψ, of the perigee from the sub-solar point were 120° and 75°, respectively. Finally, it is clear that the electromagnetic solar radiation represented by 10·7 cm was decreasing after the occurrence of its peak two or three days before. These two events

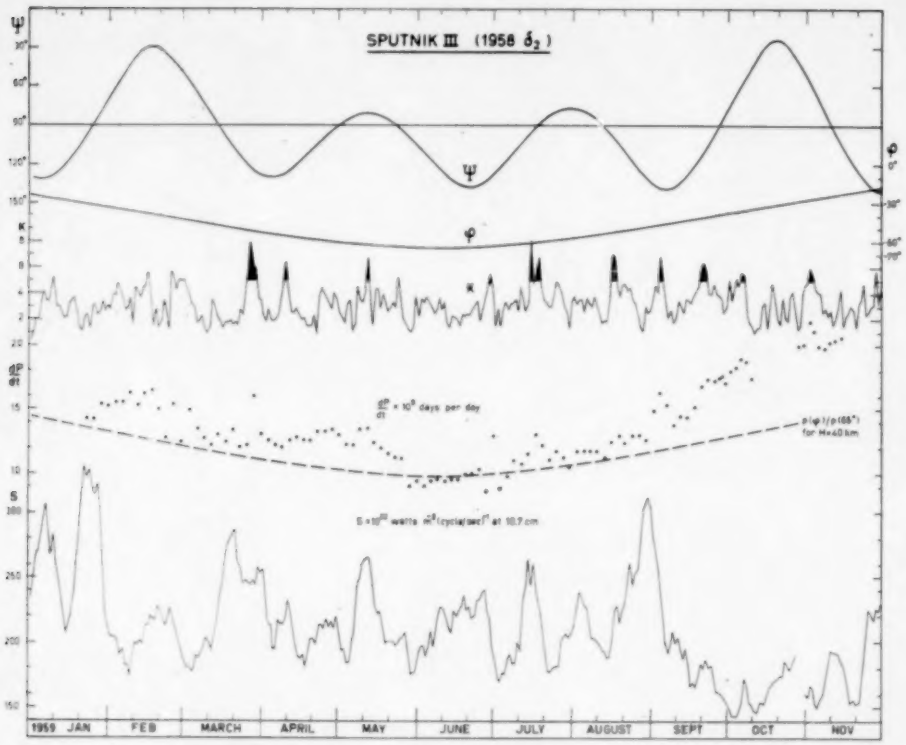


Fig. 6. Orbital acceleration of Sputnik III (1958 δ2). $-dP/dt \propto \rho H^{1/2}$ [43]. Other symbols, see Fig. 4.

show that a limited resolution corresponding to several days may smooth the transient disturbances affecting the atmospheric density.

However, if we consider the periods when the perigee was in a dark atmosphere (Ψ in Fig. 7) we can see three other increases associated with the K indices ($K > 5$) even if the resolution is only 25 revolutions. They are 28–30 June, 24–26 September and 22–24 October, when the angular distance, Ψ, of the perigee from the subsolar point is greater than 90°. Again the corresponding latitudes of the perigee are very low; between 15°N and the equator from September to October.

Taking again the Smithsonian data for Sputnik II (1957 β)[29], it can be shown (Fig. 8) that certain variations of the density can be associated with $K \geqslant 5$ indices. Two remarkable associations are found on 1–2 January and 9–11 February when the Ψ angles were 80° and 120° and the perigee latitudes were 30°N and 15°N, respectively.

Even if the observations of 1958 δ 2 are irregularly distributed three pronounced increases in the acceleration curve of Sputnik III (1958 δ 2) can be seen (Fig. 6) for 27 March, 29 June and 4 September. These are associated with geomagnetic disturbances which are represented by the K indices[2]. It should be noted that these three remarkable increases in the atmospheric density were observed during the three periods when the perigee was in the dark atmosphere. This shows again that the reactions of the atmosphere to a "corpuscular effect" are more clearly distinguishable when effects on a sunlit atmosphere can be eliminated.

If we consider periods during which the atmosphere was sunlit (May, July, August) important magnetic disturbances were also observed, but their associated effects do not

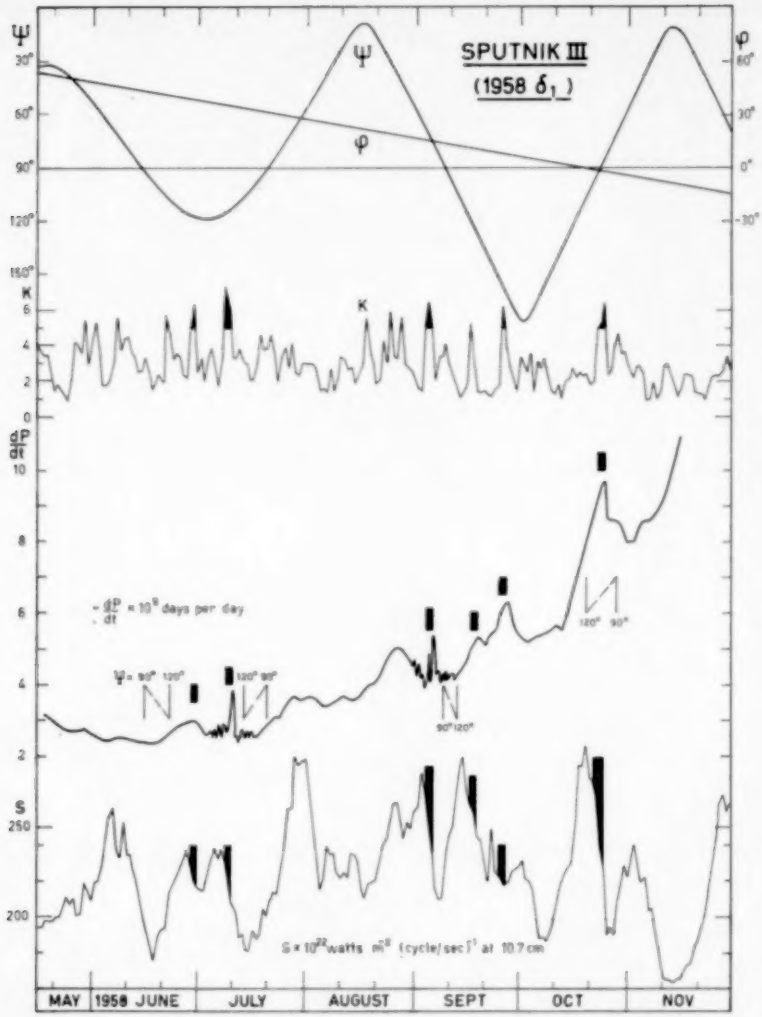


Fig. 7. Orbital acceleration of Sputnik III (1958 δ 1). $-dP/dt \propto \rho H^{1/2}$ [31]. Other symbols, see Fig. 4.

seem to be so disturbing to the behaviour of the acceleration of satellite 1958 δ 2. In particular, the magnetic storm of 12 May 1959 (see the *K* indices Fig. 6) observed by Ney *et al.*[84] in association with cosmic radiation does not show, with the low resolution of the acceleration curves, a special effect on the acceleration of the satellite. The maximum in the acceleration curve, about 8 May, is more closely associated with the maximum of the solar electromagnetic radiation. Similarly, the maximum near 15 July at sunrise appears to be related to the electromagnetic solar radiation, although the cosmic-ray bursts described by Winckler[85] are noteworthy.

Moreover, a comparison should be made between the magnetic storm of 16 August 1959 and the recurrent storm of 4 September (during

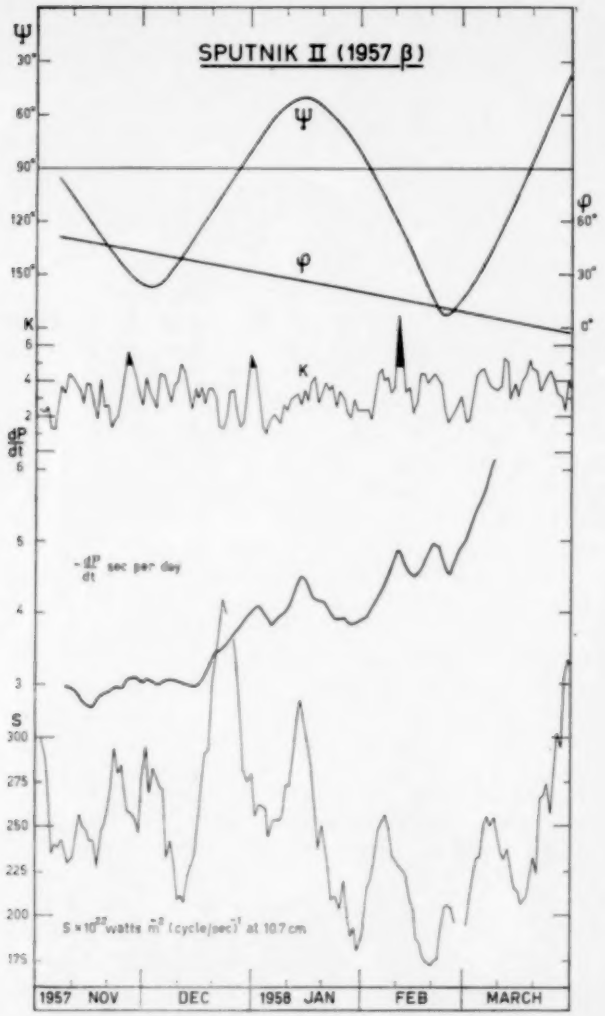


Fig. 8. Orbital acceleration of Sputnik II (1957 β). $-dP/dt \propto \rho H^{1/2}$ [29]. Other symbols, see Fig. 4.

the night for 1958 δ 2). During the first 24 hr of the magnetic storm of 16 August, Arnoldy *et al.*[1] found that about 3/4 of the particles in the outer Van Allen belt had been removed and had penetrated into the atmosphere. But the effect of the "corpuscular" heating of the atmosphere is probably not greater than the electromagnetic heating, whilst the exceptional effect observed on 4 September is clearly a transient heating in the dark atmosphere.

These different examples show that it should be possible to distinguish between all of the external effects by a detailed determination, with sufficient resolution, of the acceleration of the satellites. In any event, the diurnal variation and the 28-day periodicity show that the absorption of solar electromagnetic radiation is the primary process for heating the atmosphere above 100 km. The "corpuscular" radiation with its associated processes can easily affect the

nocturnal conditions but is more difficult to detect in a sunlit atmosphere which is already heated by an increased electromagnetic radiation. Furthermore, the strong diurnal variations at the higher altitudes can mask such effects. Finally, the penetration of energy below 100 km has practically no influence on the structure of the thermosphere in that the energy involved in flares is not important compared with the total kinetic energy of the atmosphere. The most remarkable phenomenon near 200 km is the relation with the energy of the solar radiation which varies in accordance with changes in the solar activity. This is the reason that several investigators have found correlations with various indices of solar activity. Such a variation at 200 km means that the atmosphere between 100 and 200 km is affected and must correspond to a general increase of the scale height in the entire layer.

The variations of the density as a function of altitude demonstrate that the diurnal variation is magnified with increasing height. It is the dominant factor at higher altitudes and must be associated with the gradient of the scale height and altitude of the thermopause.

The variation from one day to another is closely associated with solar activity and the size of the variations of the density, increasing with the altitude, depends on the energy of the solar electromagnetic radiation which is available during a 27-day period.

A "corpuscular effect"* is evident at the time of magnetic disturbances but it is only introduced sporadically and the energies which are involved are generally less than ultra-violet energies. Seasonal and latitude effects can only be secondary in relation to the complex effects of the diurnal and solar variations.

3. THE CONSTITUTION OF THE THERMOSPHERE

The constitution of the thermosphere, that is to say of the atmosphere above 85 km, theoretically depends on the state of molecular dissociation at the highest altitudes. Indeed, the essential observation that the scale height H increases with altitude necessitates an analysis of the variations of three parameters: the temperature T, the mean molecular mass m and the acceleration due to gravity, g.

If the atmosphere above a certain altitude was in a state of complete dissociation, an increase of H with altitude could only be explained by an increase of temperature with altitude. Above 300 km, we can neglect an important heating effect by ultra-violet radiation in that there is practically no absorption of such radiations at very high altitudes. In this case, a very pronounced dissociation of molecular oxygen below 200 km, with a low pressure (10^{-4} mm Hg) at 100 km, leads to an atomic oxygen atmosphere. To explain the gradient of the scale height it is necessary to introduce a flux of external heat transported by conduction. Such an application has been made by Nicolet[55, 57] using Chapman's theory of an extension of the solar corona.

But an atmospheric model in which the pressure at 100 km reaches 3×10^{-4} mm Hg, with a small percentage of the oxygen dissociated, represents more closely the atmospheric conditions leading to a further understanding of the thermosphere. An important temperature gradient exists below 200 km and leads to a large abundance of molecular nitrogen at high altitudes[55, 59]. Since atomic nitrogen is a secondary constituent[56, 57] we may consider an atmosphere whose constitution depends essentially on the ratio N_2/O.

In a general study, Nicolet[59] has shown how the problem of the conditions in the thermosphere can be analysed. Using as an example the following conditions at 100 km:
$p = 3 \times 10^{-4}$ mm Hg; $T = 200°$K;

$$\rho = 6{\cdot}6 \times 10^{-10} \text{ g cm}^{-3} \qquad (3.1)$$

corresponding to the concentrations (cm^{-3})
$n(O_2) = 2{\cdot}2 \times 10^{12}$, $n(N_2) = 1{\cdot}1 \times 10^{13}$;

$$n(O) = 1{\cdot}4 \times 10^{12} \text{ and } M = 27{\cdot}4 \qquad (3.2)$$

the conditions at higher altitudes can be determined if we fix the level for the beginning of diffusion and the gradient of the scale height.

In Fig. 9 are shown conditions such that diffusion begins at 120 km and at 150 km, and the scale height has large gradients of $\beta = 1{\cdot}0$,

* "Corpuscular effect" means no direct electromagnetic radiation effect.

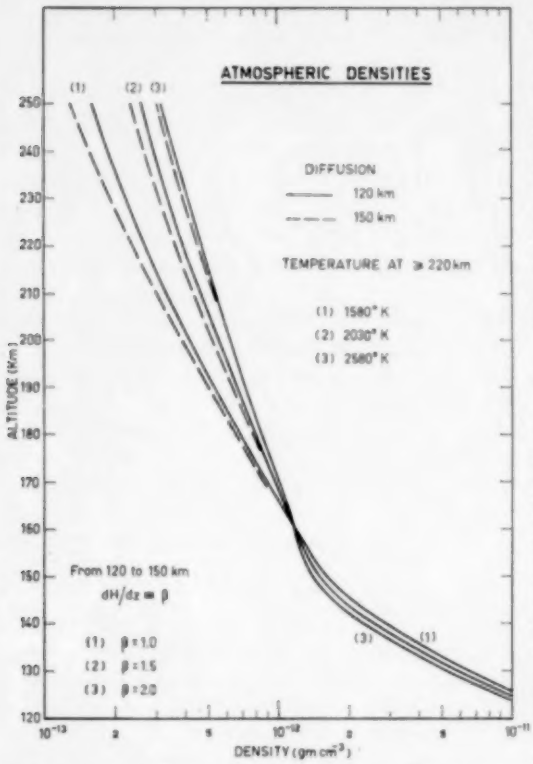


Fig. 9. Density–altitude relations above 120 km for various gradients of temperature. The height variation of the density shows that any solution for an inadequately short range of altitudes is arbitrary even if the beginning of the diffusion is introduced at 120 km or 150 km. Boundary conditions being fixed at 120 km (T and ρ constant), variation of density at 200 km is small.

1·5 or even 2·0 between 120 and 150 km and an arbitrary gradient of $\beta = 0{\cdot}2$ between 150 and 220 km. In order to simplify the calculations it was noted that these two gradients are almost equivalent to a variable gradient diminishing with height until about 300 km.

Considering Fig. 9 we see that very different conditions do not modify to any great extent the vertical distribution of the density. For example, the density at 200 km only varies by ± 25 per cent from the mean value for the above range of variables. Consequently we can conclude that, for constant conditions at 100 km, it is easily possible to obtain the densities, deduced from the satellite observations if we consider that there is a large temperature gradient between 100 and 200 km. Furthermore, the vertical distribution of the density and the absolute value at 200 km are practically independent of the exact value of the temperature gradient. An increase of the order of 50 per cent in the density at 200 km would require an increase of the order of 1000°K in the temperature. An increase of the density at 200 km should be caused mainly by its increase in the region of 120 km.

A similar effect explains the differences in the observations made with rockets at different periods at White Sands and Ft. Churchill. Similarly, differences obtained at the same place of observation can only result from a change of temperature in the lower thermosphere. In other words, the structure of the thermosphere depends primarily on the conditions at the limits applicable to the lower thermosphere (different densities at 100 km). Moreover the fact of having introduced diffusion at 120 km or 150 km does not modify the above conclusions for the atmosphere at altitudes of less than 250 km.

We can therefore conclude that: *the density of the high atmosphere is essentially dependent on the solar energy absorbed below 200 km.* This energy determines the temperature gradient up to 400 km in a sunlit atmosphere and fixes the temperature of the isothermal atmosphere up to the highest altitudes.

Since various temperature gradients between 120 km (E-layer) and 200 km (F-layer) modify only slightly the values of the density at 200 km, if constant boundary conditions are taken at 120 km, the small variations in the densities deduced from satellite observations near 200 km are easily explained. But it must be pointed out that variations must occur in the E-layer which modify the boundary conditions for the whole thermosphere. The computations lead to the following results:

$$120\ \text{km},\ \rho = 3{\cdot}26 \times 10^{-11}\ \text{g cm}^{-3};\quad T = 262^\circ\text{K} \tag{3.3}$$

$$150\ \text{km},\ \rho = (1{\cdot}5 \pm 0{\cdot}03) \times 10^{-12}\ \text{g cm}^{-3};\quad 725^\circ\text{K} \leqslant T \leqslant 1650^\circ\text{K} \tag{3.4}$$

Thus, it is certain that varying conditions occur inside the *E*-layer.

The total kinetic energy varies between 5×10^4 and 5×10^5 ergs cm^{-2} from about 120 km to 100 km. If an energy of the order of 1 erg cm^{-2} sec^{-1} is available for the heating of the *E*-layer, it is clear that the boundary conditions at 120 km will be subject to variations resulting from solar activity. In the same way, the region of the *F1*-layer where the total kinetic energy is less than 5×10^4 erg cm^{-2}, will be strongly affected by an ultra-violet heating of the order of 1 erg cm^{-2} sec^{-1}.

Below 100 km, at 85 km for example, the kinetic energy of a vertical column is of the order of 5×10^6 erg cm^{-2} and an equivalent solar energy during one day of 12 hr would require a total absorption of about 100 erg cm^{-2} sec^{-1}. We can therefore conclude that the entire thermosphere above 100 km is essentially dependent on the solar energy and its fluctuations if the energy available for heating is of the order of 1 erg cm^{-2} sec^{-1}. Such an energy input must be found in the X-ray and ultra-violet radiations absorbed above 100 km.

4. SOLAR RADIATION

The diurnal variation of the atmospheric density above 250 km shows clearly that the heating of the thermosphere depends mainly on the electromagnetic radiation absorbed in the atmosphere above 100 km.

Above 100 km the atmosphere absorbs radiations with wavelengths of less than 1750 Å as a result of the dissociation of the oxygen molecules. The recombination of oxygen atoms mainly occurs below 100 km after a downward transport and for this reason there is no important heating available from *in situ* recombinations. However, the energy available from the various monochromatic radiations between 1500 Å and 1300Å corresponds to about 3 ergs cm^{-2} sec^{-1}, and the part transformed immediately into heat is about 0·5 erg cm^{-2} sec^{-1} in the *E*-layer. Furthermore, the total energy from X-rays absorbed in the *E*-layer does not exceed 0·5 erg cm^{-2} sec^{-1} at the maximum of the solar cycle. In addition, the heat flow from upper levels will lead to an important effect in the *E*-layer since convection and conduction are certainly involved. We can conclude therefore that the *E*-layer is subject to varying conditions depending on the coronal flux in the X-ray spectrum and on the chromospheric and coronal flux in the ultra-violet spectrum.

At altitudes above the *E*-layer, from the beginning of the *F1*-layer, we must take into account the absorption of radiations of wavelengths longer than 200 Å. This corresponds to the entire ultra-violet spectrum involving helium lines at 304 Å and 584 Å and also coronal lines of highly ionized atoms. An energy of the order of 1 erg cm^{-2} sec^{-1} can produce temperature gradients of the order of 20°K per km. Such important gradients lead to very high temperatures above 200 km.

If we consider a steady state for an overhead sun, it is easy to show (see Section 5) that the energy E is distributed with height as follows

$$E = E_\infty + E_{uv}(1 - e^{-\tau}) \tag{4.1}$$

in which E_∞ denotes the energy flow at the top of the layer, E_{uv} the ultra-violet solar energy available at the top of the earth's atmosphere for a certain spectral range and τ is the optical depth, which is defined by

$$\tau = K_\lambda \int_z^\infty n\,dz \tag{4.2}$$

Since the heating must be proportional to the energy absorbed, it is evident that the temperature gradient will decrease with height up to a certain altitude where the optical depth is negligible.

If there is no external heat flow, the sunlit atmosphere must be isothermal where there is no absorption of solar radiation. Since X-rays and ultra-violet radiations are absorbed between the *E*-layer and the *F*-layer the temperature must increase continuously and no permanent decrease can be expected in the whole thermosphere.

In order to study the variation of the short ultra-violet radiations and of X-rays one refers to measurements of the centimeter or decimeter electromagnetic radiations emitted by the solar corona. Nicolet[69] has compared the results

obtained at various observatories from 3 to 30 cm and has shown that the general behavior of the solar activity is the same for the entire spectral region. The mean daily solar fluxes from 8 to 15 cm are subject to identical variations. The extremes of 3 cm and 30 cm differ from 10 cm by an extent which is much less than the fluctuations. In other words, the largest fluctuations of the solar radiation from 3 cm to 30 cm are observed at about 10 cm although the complete spectrum varies, under the same conditions, with the solar activity.

Since Priester and Martin[69], and Jacchia[33] after having used 10 cm[32], have based their analysis of satellite accelerations on the solar energy of 20 cm wavelength observed at Berlin, it is important to compare the observational data at 10 cm and 20 cm. Fig. 10 shows that, in 1958, the solar radiation at 10·7 cm measured at Ottawa does not indicate a general variation which is as pronounced as that for the solar radiation at 20 cm measured at Berlin. In fact, during 1958, the relation between the extremes of radiation at 20 cm is of the order of 5 while that for 10·7 cm (which must be a maximum) is only about 2. Since Priester and Martin[69] and Jacchia[33] are able to follow the general variation of the atmospheric densities deduced from satellite measurements by using the 1958 measurements at 20 cm, it is quite strange that such a good correlation has been obtained. There is no way to account for a special behavior at 20 cm, with a variation of a factor of 5 in 1958, when the entire spectrum between 3 cm and 30 cm does not vary by more than a factor of 2.

Fig. 11 shows that, in 1959, the solar radiations measured at 10·7 cm in Ottawa and at 20 cm in Berlin are subject to the same general variations and that the ratio Berlin/Ottawa of the solar fluxes, taking as a unit the mean value for 1959, does not vary by more than ±20 per

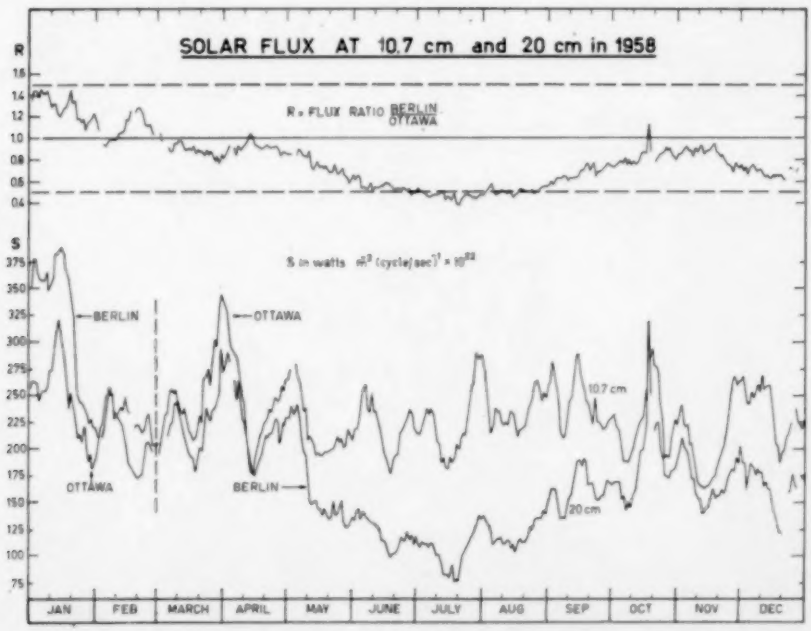


Fig. 10. Comparison between solar fluxes at 10·7 cm (Ottawa) and at 20 cm (Berlin) in 1958. While the difference between maximum and minimum corresponds to a factor of the order of 5 at Berlin, it is only of the order of a factor 2 at Ottawa. The real variation cannot be more than a factor of 2 and the flux ratio of 20 cm/10 cm cannot vary from 1·5 to 0·4.

cent. The data on radio radiations at 20 cm, therefore, introduce a systematic error when used in 1958 to compare the variations of the atmospheric density with that of the solar flux.

In conclusion, when errors in calibration of radio measurements have been eliminated one can say that in general the variations of the activity are represented by all wavelengths between 3 and 30 cm. The most pronounced range of the variations is given by the spectral region centered on 10 cm. Thus the observations at Ottawa at 10·7 cm, made during 1958 and 1959, and indicating a maximum variation of a factor 2·2, fix the maximum of the variation of the mean daily value for all the spectral region from 3 cm to 30 cm, i.e. of the slowly varying component of the sun's radio-emission.

Since the temperature and its vertical distribution in the chromosphere and corona remain normal in the active regions emitting the decimeter radiations, it is convenient to consider as the basic radiation of the sun heating the upper atmosphere the minimum value of the radio flux. Hence we may assume that from July to November 1958 the basic radiation S at 10·7 cm was never less than 220 units [watts m^{-2} (cycle-sec)$^{-1} \times 10^{-22}$], while in 1959, from the beginning of September to the end of November, it was less than 220 reaching 150 at the beginning of September.

Such differences must correspond to variations of the solar emission in the ultra-violet region and particularly in the X-ray spectrum. It is evident that the variation in amplitude of the radio emission cannot be directly proportional to that of the entire ultra-violet spectrum or of the spectrum at the shortest wavelengths. The difficulty of obtaining an exact relation between the optical and radio ranges is easy to understand since such radiations originate between the photosphere and corona, i.e. from the lower and cooler part of the solar atmos-

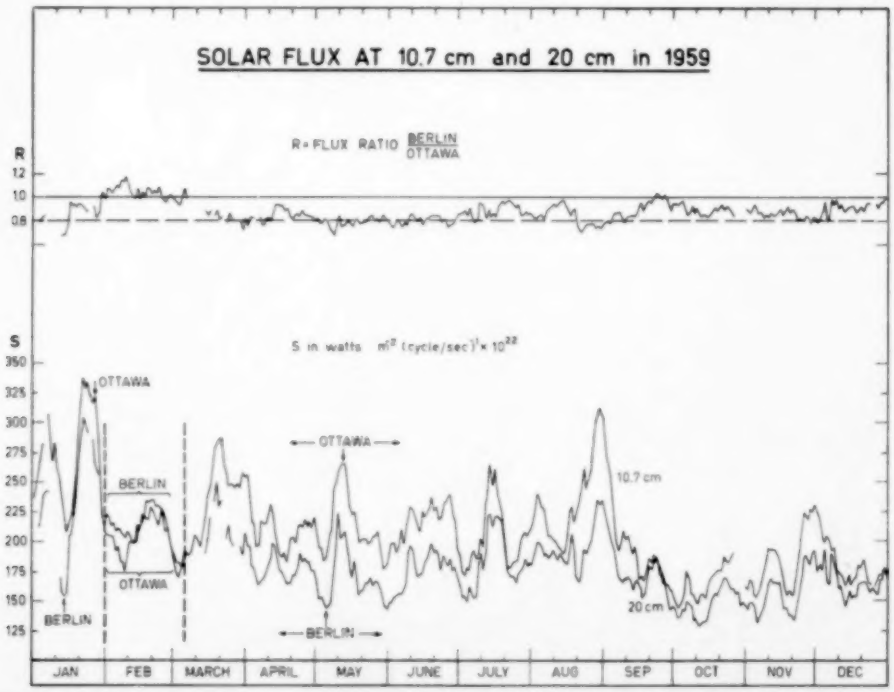


Fig. 11. Comparison between solar fluxes at 10·7 cm (Ottawa) and at 20 cm (Berlin) in 1959. The flux ratio 20 cm/10 cm varies only about ±20 per cent and such a small variation indicates that 1958 results on 20 cm cannot be accepted.

phere up to the normal high temperature corona. Nevertheless, it is true that the terrestrial atmosphere between 100 km and 200 km is subject to the variable heating resulting from the absorption of ultra-violet radiation and X-rays. The implication is that the thermosphere is heated by all radiations for which the absorption cross-section is greater than 10^{-19} cm^2. There is no particular terrestrial layer above 100 km which is especially heated, since the energy is distributed with height according to (4.1).

Sources of heating other than electromagnetic radiation depend on the energy which is available. The use of the energy of the radiation belt in one form or another is limited by the total energy available. This, according to Dessler and Vestine[15], is 6×10^{22} erg and is about a factor of 10 greater than the energy involved in a magnetic storm; 10^{22} ergs according to Chapman and Bartels[10]. On the other hand, Arnoldy *et al.*[1] have observed, during the first phase of a geomagnetic storm, a loss of about two-thirds of the energy of the outer belt. Such energies place limits on the energy which would be immediately available from the radiation belt and we agree with Bates[4] that the time period would be too short if "corpuscular" energy must pass through the radiation belt.

With a total of 6×10^{22} ergs, the maximum energy available for the earth's atmosphere cannot be more than 10^4 erg cm^{-2}. An energy supplied to the atmosphere of the order of 1 erg cm^{-2} sec^{-1}, which would be given by electrons of the order of 10 keV according to Krassovsky[44], will lead to the total energy of the radiation belt in about 3 hr. The hypothesis put forward by Krassovski[44] is very difficult to accept since the energy of the radiation should be renewed several times a day. Consequently, various suggestions of the normal heating of the upper atmosphere or of the creation of a heated atmosphere in a definite region of the auroral zone by the channelling of charged particles through the radiation belt cannot be accepted. Hence the deduction of Jastrow[35] that two atmospheres can exist simultaneously, i.e. a low latitude atmosphere with a temperature of the order of about 1000°K and one at a high latitude with a temperature of about 2000°K, has little plausibility.

An inflow of 2×10^{11} electrons cm^{-2} sec^{-1} leading to a heat source $Q=4\times10^{16}$ cal cm^{-3} sec^{-1} at 300 km, as computed by Jastrow[35], corresponds to an injection into the thermosphere which is not observed even during auroras. Meredith[50] has given a specific example of a flux between 2 and 6×10^7 electrons cm^{-2} sec^{-1} of about 10 keV inside a diffuse aurora. Such an example shows that *normal* heating of the atmosphere by corpuscular radiation must be excluded. However, sporadic conditions corresponding to geomagnetic storms and leading to energies of 10^{22} ergs may lead to appreciable transient heating since it corresponds to 0·5 erg cm^{-2} sec^{-1} for the entire earth during one hour or about one-twentieth of the earth's surface during 24 hours.

Dessler[14] has put forward the view that the normal heating of the thermosphere is due to hydromagnetic heating. Against this view it can be stated that the diurnal variation of the temperature of the thermosphere cannot be explained by this process. But the increased orbital acceleration of satellites observed during geomagnetic storms at any latitude does not exclude a direct heating of the atmosphere outside the auroral zone. The hydromagnetic heating, therefore, must be kept among the hypotheses needed to explain a heating at very low latitudes during geomagnetic storms. A final decision cannot be reached until the variations of the atmospheric density are properly explored by using physical parameters such as the temperature and pressure in addition to the density and the scale height.

5. TRANSPORT OF HEAT BY CONDUCTION

5.1. *The equations of conduction*

The solar radiation heats the thermosphere at various altitudes, according to the values of the coefficients of absorption which range from 10^{-19} cm^2 to 10^{-16} cm^2. It is, therefore, appropriate to investigate the behavior of the atmosphere under the effect of the conduction of heat depending on the gradient of temperature.

The flux density of heat, E, can be written

$$E = -\lambda_e \operatorname{grad} T \tag{5.1}$$

where T is the temperature and λ_e is the thermal conductivity. If E is expressed in erg cm^{-2} sec^{-1}, λ_e is given in erg cm^{-1} sec^{-1} deg^{-1} and is related to the coefficient of viscosity, μ, by the equation[11]

$$\lambda_e = f \mu c_v \tag{5.2}$$

where c_v is the specific heat at constant volume and f represents a numerical value equal to about 2·5 for spherical molecules (monatomic gas) and can be equal to 1·9 for diatomic molecules.

The viscosity μ can be expressed as

$$\mu = \frac{5}{16} \frac{(\pi k m T)^{1/2}}{\pi \sigma^2} \tag{5.3}$$

where σ is the atomic radius, k Boltzmann's constant and m the atomic mass. This formula can be used in the high atmosphere[59, 60], and the values used are, for atomic hydrogen, $\sigma = 2{\cdot}0 \times 10^{-8}$ cm,

$$\mu(H) = 6{\cdot}8 \times 10^{-6} T^{1/2} \tag{5.4}$$

atomic oxygen, $\sigma = 2{\cdot}4 \times 10^{-8}$ cm,

$$\mu(O) = 1{\cdot}9 \times 10^{-5} T^{1/2} \tag{5.5}$$

molecular nitrogen and oxygen, $\sigma = 3{\cdot}3 \times 10^{-8}$ cm

$$\mu(N_2, O_2) = 1{\cdot}3 \times 10^{-5} T^{1/2} \tag{5.6}$$

In using equations (5.2) and (5.3) the thermal conductivity can be denoted by

$$\lambda_e = A T^{1/2} \tag{5.7}$$

where A is a constant, depending on the atmosphere constituents. The numerical value of this constant is (in erg cm^{-1} sec^{-1} $deg^{-3/2}$):

$$A(H) = 2{\cdot}1 \times 10^{3} \tag{5.8}$$

$$A(O) = 3{\cdot}6 \times 10^{2} \tag{5.9}$$

$$A(O_2, N_2) = 1{\cdot}8 \times 10^{2} \tag{5.10}$$

Since λ_e is a function of the temperature, we introduce a new variable defined by

$$\theta = \int_{T_2}^{T} \frac{\lambda}{\lambda_2} dT \tag{5.11}$$

leading to the following relation, by using (5.7),

$$\theta = \frac{2}{3} \frac{T^{3/2} - T_2^{3/2}}{T_2^{1/2}} \tag{5.12}$$

Using (5.7) and (5.12) equation (5.1) can now be written as

$$E = -A T_2^{1/2} \operatorname{grad} \theta \tag{5.13}$$

The continuity equation is written as follows

$$\rho c_v \frac{\partial T}{\partial t} + \operatorname{div} E = P - L \tag{5.14}$$

in which P denotes the production of heat per unit of time, per unit of volume and L the loss of heat per unit of time and volume. Considering the heat capacity per unit of volume, ρc_v, where ρ is the density, equation (5.14) can be written as

$$\frac{\partial \theta}{\partial t} = \frac{A T^{1/2}}{\rho c_v} \nabla^2 \theta + (P - L) \frac{T^{1/2}}{\rho c_v T_2^{1/2}} \tag{5.15}$$

where the coefficient $A T^{1/2} / \rho c_v$ is the thermal diffusivity which, if n is the atomic concentration, can be written

$$\alpha(T, n) = A_1 T^{1/2} / n \tag{5.16}$$

where A_1 is a constant which has the following values (cm^{-1} sec^{-1} $deg^{-1/2}$)

$$A_1(H) = 1{\cdot}0 \times 10^{19} \tag{5.17}$$

$$A_1(O) = 1{\cdot}75 \times 10^{18} \tag{5.18}$$

$$A_1(N_2, O_2) = 5{\cdot}3 \times 10^{17} \tag{5.19}$$

As a result, the differential equation (5.15) for the conduction becomes:

$$\frac{\partial \theta}{\partial t} = \frac{A_1 T^{1/2}}{n} \nabla^2 \theta + \frac{A_1 T^{1/2}}{n A T_2^{1/2}} (P - L) \tag{5.20}$$

By using the variable T instead of introducing the temperature parameter θ, equation (5.20) would be non-linear. In fact the problem can always be studied by considering a condition at the limits $\theta_2 = 0$ for all fixed values of $T = T_2$.

If $\partial\theta / \partial t = 0$, equation (5.20) becomes Poisson's equation which can be applied to the steady state

$$\nabla^2 \theta + \frac{P - L}{A T_2^{1/2}} = 0 \tag{5.21}$$

If, on the other hand, there is no loss or production of heat at the centre of the volume being studied then (5.20) corresponds to the Laplace equation

$$\nabla^2\theta = 0 \tag{5.22}$$

When there is cooling by conduction with no production or loss of heat inside of the volume, equation (5.20) should be written

$$\frac{\partial\theta}{\partial t} = \frac{A_1 T^{1/2}}{n}\nabla^2\theta \tag{5.23}$$

in which, for ease of calculation, we can take a mean value of $A_1T^{1/2}$ and, thus, the thermal conduction depends, essentially, upon the value of the concentration n.

5.2. *Steady state with no source or loss of heat inside the volume*

In a sphere where the temperature is a function only of the radius r, the solution of (5.20) is

$$\theta = \theta_1 \frac{r - r_2}{r_1 - r_2} \times \frac{r_1}{r} \tag{5.24}$$

if $\theta_2 = 0$ at a distance $r = r_2$ from the center of the sphere and θ_1 at $r = r_1$. According to (5.12), the distribution of the temperature becomes, (5.24),

$$(T^{3/2} - T_2^{3/2})/(T_1^{3/2} - T_2^{3/2}) = = \frac{r - r_2}{r_1 - r_2}\frac{r_1}{r} \tag{5.25}$$

The heat flow E_r is therefore (5.13) and (5.25),

$$E_r = \frac{r_2}{r}\frac{2}{3}A\,A\,(T^{3/2} - T_2^{3/2})/(r - r_2) \tag{5.26}$$

An immediate application is a night-time atmosphere where there is no heating by ultra-violet radiation at a sufficiently high altitude. In an atomic oxygen atmosphere if the heat flow has the following values at 500 km

0·1 0·2 0·5 erg cm^{-2} sec^{-1},

the corresponding values of the temperatures obtained at 700 km are

1250°K; $\Delta T = 350$ 1550°K; $\Delta T = 650$
2300°K; $\Delta T = 1400$

1850°K; $\Delta T = 250$ 2100°K; $\Delta T = 500$
2900°K; $\Delta T = 1300$

if the temperatures T_2 at 300 km are 900°K and 1600°K, respectively.

These numerical data show how the temperature would vary if an external heating were involved. However, since no direct determination of the temperature can be obtained from satellite data, it is necessary to consider (5.1) when the scale height H is used instead of T. With (5.9) and (5.10), the equations are

$$E\,(\mathrm{O}) = -0{\cdot}817 \times 10^{-3}\,(g/900)^{3/2}H^{1/2}\frac{dH}{dz} \tag{5.27}$$

for atomic oxygen, and

$$E\,(\mathrm{N_2}, \mathrm{O_2}) = -0{\cdot}945 \times 10^{-3}\,(g/900)^{3/2}H^{1/2}\frac{dH}{dz} \tag{5.28}$$

for air, when $g = 900$ cm sec^{-1} is adopted for an altitude of 280 km.

If we consider the night-time data of Jacchia[33] between 600 and 700 km, it is clear that the energy which is required to maintain such a gradient according to (5.27) and (5.28), is

$$E\,(600\text{–}700\ \mathrm{km})_{\mathrm{night}} = (0{\cdot}28 \pm 0{\cdot}02)\ \mathrm{erg\ cm^{-2}\ sec^{-1}} \tag{5.29}$$

However, it has been shown (see Fig. 2) that at the *highest levels* it is possible to interpret the satellite observations concerned with atmospheric drag in terms of an isothermal atmosphere subject to diffusion.

As far as the daytime data are concerned, Jacchia[33] has deduced a more pronounced gradient which leads, in the same range of altitudes, to

$$E\,(600\text{–}700\ \mathrm{km})_{\mathrm{day}} \geqslant 1\ \mathrm{erg\ cm^{-2}\ sec^{-1}} \tag{5.30}$$

As a matter of fact, such a deduction corresponds to an increase of E with altitude, since near 600 km $E \simeq 1$ erg cm^{-2} sec^{-1} and near 700 km $E > 1{\cdot}3$ erg cm^{-2} sec^{-1}.

If at 700 km, a difference of about 1 erg cm^{-2} sec^{-1} distinguishes the difference between a sunlit and dark atmosphere (5.29) and (5.30), it would be necessary to assume that electromagnetic radiation can be absorbed at such high altitudes. The maximum possible total number of atoms in a vertical column, in using Jacchia's data[33], cannot be more than 8×10^{14} cm^{-2}.

Since any absorption cross-section cannot be more than 10^{-16} cm^2, the optical depth cannot be more than 0·08 and the ultra-violet energy needed would be more than 12·5 erg cm^{-2} sec^{-1}. Furthermore, since the absorption cross-section must be less than 10^{-16} cm^2, the ultra-violet radiation required should be more than 50 erg cm^{-2} sec^{-1} and such an energy will lead to temperature gradients below the unit optical depth of more than 500°K per km.

In fact, the gradients of temperature for $400°K \leqslant T \leqslant 1600°K$ are given by (E in erg cm^{-2} sec^{-1}):

$$\left(\frac{dT}{dz}\right)_{km} = (10 \pm 3)\, E \qquad (5.31)$$

for an atomic oxygen atmosphere, and

$$\left(\frac{dT}{dz}\right)_{km} = (20 \pm 6)\, E \qquad (5.32)$$

for an undissociated atmosphere. The scale height gradients, β, are under the same conditions, respectively

$$\beta\,(O) = (0{\cdot}51 \pm 0{\cdot}10)\, E \qquad (5.33)$$

and

$$\beta\,(N_2, O_2) = (0{\cdot}46 \pm 0{\cdot}10)\, E \qquad (5.34)$$

It is clear that E cannot be very different between day and night at highest altitudes since it would require a higher ultra-violet energy than is available from the sun. A corpuscular effect will not lead to a difference between day and night, and any external heating cannot explain such a difference. It must be concluded, therefore, that there is no physical process to be found to explain an increase of the temperature at such altitudes. An increase of the scale height gradient can be found only where the laws of diffusion can be employed, in fact at very high altitudes the gradient decreases and finally the scale height becomes constant.

Another aspect of a special behavior deduced from observational data is the gradient of the scale height (see Fig. 3) adopted by King-Hele and Walker[42] near 200 km. The gradient is negative between 210–220 km, $\beta = -0{\cdot}3$; it is positive between 200–210 km, $\beta = +0{\cdot}4$. Using formulas (5.27) and (5.28), a thin layer near 200 km should lead to an upward heat flow of $(0{\cdot}80 \pm 0{\cdot}06)$ erg cm^{-2} sec^{-1} and a downward heat flow of $(1{\cdot}15 \pm 0{\cdot}08)$ erg cm^{-2} sec^{-1}, i.e. a heat loss of about 2 erg cm^{-2} sec^{-1} for a layer of thickness less than 10 km. The total energy lost in one hour by such a layer would be more than its total kinetic energy. This shows that such a discontinuity cannot be a permanent feature in the thermosphere. We shall see later that such an artificial gradient would have a very short lifetime. In any case, it can be shown (see Fig. 3) that the vertical distribution of the density can be represented in an isothermal atmosphere subject to diffusion without such an anomaly in the scale height gradient. Therefore, the vertical distribution of the density near 200 km is not very sensitive to the exact value of the scale height gradient.

5.3. *Steady state with heating by ultra-violet radiation and loss by infra-red radiation*

In a steady state with one dimension the equation (5.21) is

$$\frac{\partial^2 \theta}{\partial z^2} + \frac{P - L}{A T_2^{1/2}} = 0 \qquad (5.35)$$

or

$$\frac{\partial E}{\partial z} + P = L \qquad (5.36)$$

Considering an atmosphere subject to a heating by several radiations, the production of heat P (cm^{-3} sec^{-1}) is

$$P = \Sigma n K E_{uv} \exp\left(-\int_z^\infty n K dz\right) dz = \Sigma E_{uv} e^{-\tau} d\tau \qquad (5.37)$$

in which K is the appropriate absorption cross-section of the radiation E_{uv} of wavelength λ by atoms of concentration n. τ is the optical depth defined by

$$d\tau = -nK dz \qquad (5.38)$$

leading to

$$\tau = \int_z^\infty n K dz \qquad (5.39)$$

In the thermosphere, the principal infra-red radiator is atomic oxygen, according to Bates's process[3]

$$O\,(^3P_1) \longrightarrow O\,(^3P_2) + h\nu\ (\lambda = 63\mu) \qquad (5.40)$$

since molecular nitrogen and oxygen have no dipole. The loss of heat (cm^{-3} sec^{-1}) is therefore (in erg cm^{-3} sec^{-1})

$$L = Rn(O) = n(O)\,(1{\cdot}68 \times 10^{-18} e^{-228/T})/(1 + 0{\cdot}6 \times e^{-228/T} + 0{\cdot}2 \times e^{-325{\cdot}3/T}) \qquad (5.41)$$

leading to

$$L(E\text{-layer}) = (5 \pm 1) \times 10^{-19}\, n(O) \text{ erg cm}^{-3} \text{ sec}^{-1} \qquad (5.42a)$$

$$L(F\text{-layer}) = (8 \pm 0{\cdot}5) \times 10^{-19}\, n(O) \text{ erg cm}^{-3} \text{ sec}^{-1} \qquad (5.42b)$$

Using (5.37) and (5.40), integration of (5.36) leads to

$$E = E_\infty + E_{uv}(1 - e^{-\tau}) + \bar{R} n(O) H(O) (g/\bar{g}) \qquad (5.43)$$

in which E_∞ denotes the external heat available for the layer, and $\bar{R}$ and $\bar{g}$ are mean values.

Since the absorption cross-section in the ultra-violet range is $K \geqslant 10^{-17}$ cm^2 and $R < 10^{-18}$ erg sec^{-1},

$$E_{uv}(1 - e^{-\tau}) > \bar{R} n(O) H(O) \qquad (5.44)$$

and above the unit optical depth, it is clear that the radiation loss is also negligible. A loss of heat of the order of 1 erg cm^{-2} sec^{-1} requires 2×10^{18} cm^{-2} atoms of oxygen, i.e. a layer of 10 km with concentration of 2×10^{12} atoms cm^{-3} which corresponds to an altitude of the order of 100 km. It can be concluded that the loss of heat at 63μ occurs mainly inside the E-layer, where the radiation is certainly of the order of 1 erg cm^{-2} sec^{-1}. In the F-layer the total loss of a vertical column would be of the order of 0·1 erg cm^{-2} sec^{-1}: thus the loss of heat by radiation may be neglected in a sunlit atmosphere above the E-layer when the ultra-violet energy available is at least 1 erg cm^{-2} sec^{-1}. On the other hand, any ultra-violet energy of the order of 0·1 erg cm^{-2} sec^{-1} absorbed above the level of the $F1$ peak is of the order of the heat loss by atomic oxygen since $K \simeq 10R$.

The integration of (5.43) with the use of (5.27) and (5.28) may lead to a determination of the possible increases of the scale height (H_∞) up to the isothermal layer. The result can be written as follows (neglecting the variation of g):

$$H_\infty^{1/2} = H_0^{1/2} + \frac{E_\infty}{2A} \log \tau_0/\tau + \frac{E_{uv}}{2A}\,[\log \tau_0 - Ei(-\tau_0) + 0{\cdot}57722] - \frac{\bar{R}}{2A}\,[n_0(O) H_0(O) - n(O) H(O)] \qquad (5.45)$$

where A is the constant of (5.27) or (5.28).

Adopting $\tau_0 = 1$, near the absorption peak, the effect of the ultra-violet radiation on the scale height can be shown by the following examples:

If $H_0 = 40$ km ($\tau = 1$), H_∞ has the values (± 7 per cent)

H = 60 km 84 km 113 km 145 km

when

E_{uv} = 1 2 3 4 erg cm^{-2} sec^{-1}

If $H_0 = 50$ km ($\tau = 1$), H_∞ reaches the values

H = 72 km 99 km 129 km 164 km

when

E_{uv} = 1 2 3 4 erg cm^{-2} sec^{-1}

In other words, the scale height varies by a factor of 2 between the absorption peak and the beginning of the isothermal layer when the ultra-violet energy available for heating is of the order of 2 erg cm^{-2} sec^{-1}. Since diffusion is also involved, it is evident that scale heights greater than 100 km can be reached if the ultra-violet energy is of the order of 1 erg cm^{-2} sec^{-1}.

Evidence favouring an important gradient of temperature in a sunlit atmosphere is shown by the occurrence of pronounced diurnal variations of the density. An exact vertical distribution of the temperature cannot be obtained since the solar spectrum in the ultra-violet must be known in all the details and also the exact, varying, ratio N_2/O. However, the diminution in the temperature gradient should be indicated by a law of the form

$$\frac{dT}{dz} \propto \Sigma E_{uv}(1 - e^{-\tau}) \qquad (5.46)$$

for an overhead sun.

It is possible to obtain some idea of the conditions which are required to reach the observed densities at 200 km and scale heights

deduced from (5.45). If we assume, for example, that below the specific level of 150 km (see equation 3.4), the effect of convection is such that the energy E follows the following law

$$E = E_{uv}\left(\frac{H}{H_h}\right)^{1/2} \propto \frac{v}{v_h} \tag{5.47}$$

where H_h corresponds to 150 km, $H < H_h$ and v is the kinetic energy. The scale height gradient is, according to (5.28),

$$\beta = E_h / AH_h^{1/2} = 1{\cdot}058 \times 10^3 \, (E_h / H_h^{1/2}) \tag{5.48}$$

As was said before (see equation 3.4), the density remains almost constant at 150 km for any gradient $0{\cdot}50 \leqslant \beta \leqslant 1{\cdot}50$. The energy necessary for such gradients varies according to (5.48) from 0·7 erg cm^{-2} sec^{-1} to 3·3 erg cm^{-2} sec^{-1} (see Fig. 12). The solar energy which must be available for the heating cannot be determined if the absorption cross-section is not known. Let us assume for example a mean value of 2×10^{-17} cm^2 which corresponds to the most important part of the ultra-violet spectrum. In such a case, the following ultra-violet energies

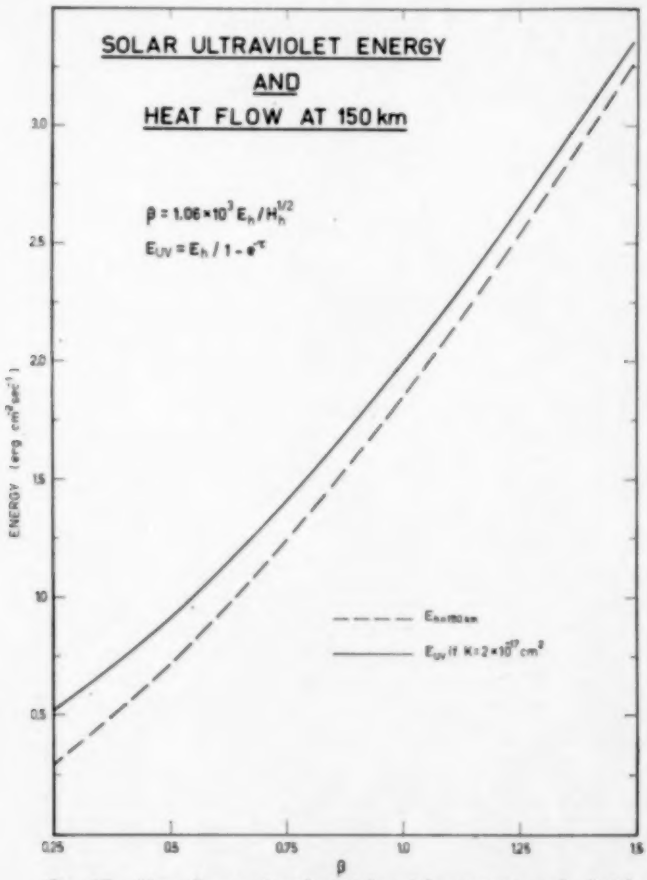


Fig. 12. Heat flow and scale height with constant scale height gradient. The effect of thermal conductivity at 150 km is shown by a curve relating the scale height gradient and the energy necessary to maintain a certain gradient. Assuming an absorption cross-section of the order of 2×10^{-17} cm^2 the ultra-violet energy necessary at the top of earth's atmosphere E_{uv} is deduced.

E_{uv} at the top of the earth's atmosphere would be necessary:

β	0·5	0·75	1·00	1·25	1·50
E_{uv}	0·9	1·4	2·0	2·7	3·4 erg cm^{-2} sec^{-1}

and the corresponding temperatures at 150 km (see Fig. 13) would be:

T	725	960	1190	1425	1650°K

Such scale height gradients and temperatures would correspond to real parameters since the total kinetic energies which are involved in the atmosphere above 150 km are between 10^4 erg cm^{-2} and 6×10^4 erg cm^{-2} corresponding to E_{uv} from 0·9 erg cm^{-2} sec^{-1} to 3·4 erg cm^{-2} sec^{-1}, respectively. In other words, the solar energy which is used during one day of 12 hr is of the same order as the kinetic energy of the vertical column above 150 km. It is, consequently, certain that the variations in the thermosphere above 100 km can be associated with the variations of the solar energy and that the variations of the temperature and its gradient are closely associated with the absorption processes below 200 km. The temperature at the thermopause depends strongly on its gradient below 200 km. But the level of the thermopause must be subject to a diurnal variation according to the laws of cooling after sunset.

5.4. *Cooling of the atmosphere after sunset*

The application of equation (5.23), after sunset, in the following form

$$\frac{\partial \theta}{\partial t} = (A_1 T^{1/2}/n) \times (\partial^2 \theta / \partial z^2) \qquad (5.47)$$

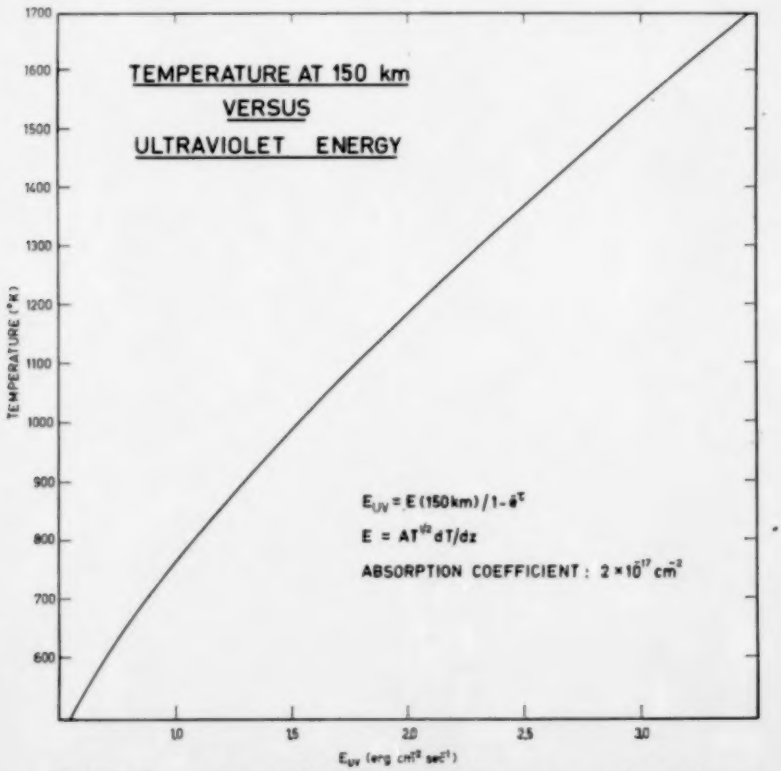


Fig. 13. Temperature at 150 km and ultra-violet energy available at the top of the earth's atmosphere. From the abscissae representing the ultra-violet energy, may be deduced the temperature at 150 km if the absorption cross-section is assumed.

corresponding to an initial atmosphere with a certain vertical gradient leads to the following results.

Let us consider, as an example, an atmosphere* giving approximately the vertical distribution of the observed density and corresponding to a heat flow of about 0·3 erg cm^{-2} sec^{-1} through atomic oxygen. If the variation is very small at 200 km, it is of the order of a factor of 10 near 650 km in about 12 hr. There is a clear indication that the atmosphere becomes rapidly isothermal at the highest altitudes and the isothermy extends with time to lower levels. It must be pointed out that $\rho H^{1/2}$ at highest altitudes decreases rapidly during the first hours, and it is interesting to note that the variation is very small between 12 and 24 hr compared with the variation during the first 12 hr.

In fact, the tendency to isothermy is very rapid when the density is less than 5×10^{-15} g cm^{-3} or concentrations less than 10^8 cm^{-3}. Equation (5.47) shows that if a temperature gradient corresponding to a heat flow of the order of 0·3 erg cm^{-2} sec^{-1} exists in an atomic oxygen atmosphere, it disappears in less than 30 min between 450 km and 750 km. Such a result shows that it is not possible to consider a dark atmosphere with a temperature gradient above 300–350 km. Moreover, the temperature of the isothermal atmosphere decreases in a continuous way following a decrease in the temperature gradient between 200 km and 300 km.

The absolute value of the decrease, being obviously a function of the initial gradient, the behavior of the variation depends upon the vertical distribution of the ultra-violet radiation absorbed in the sunlit atmosphere. Nevertheless, it is clear that the variation of the temperature of the isothermal layer between $t=30$ min and $t=2$ hr is of the same order as that between $t=2$ hr and $t=12$ hr. In other words, after a rapid tendency towards isothermy at highest altitudes, the temperature of the isothermal layer decreases according to the rapidity with which the isothermal layer extends downwards and the variation of the temperature below 200 km.

As a consequence, the heat transport by conduction associated with the gradient of temperature that results from the absorption of solar ultra-violet radiation explains the diurnal variation of the density of the thermosphere deduced from the variation of the acceleration of the satellites. We must also conclude that the diurnal variation of the temperature of the isothermal region is of the order of 500°K and that the temperature gradient is certainly subject to a variation down to 150 km.

Lastly, we must realize that the heating during magnetic storms will have an effect during a time of the order of one day since the effect of the solar electromagnetic radiation leads to a strong diurnal variation. Consequently, any sporadic effect, if energetic enough, will modify the atmospheric structure only during its own lifetime and its role will be more or less effective if it occurs in a dark or sunlit atmosphere, respectively.

5.5. *Times of conduction*

The application of equation (5.47) to the problem of the cooling of the atmosphere after sunset indicates that the time required for the transport of heat varies greatly with altitude and distance. Actually, the time of conduction is proportional to the concentration, n, and to the square of the distance.

In order to provide order of magnitudes for times of conduction, we apply equation (5.47) to very simple examples. Let us consider two infinite regions of the atmosphere where the conditions are such that the temperatures are initially T_1 ($x \leqslant 0$) and T_2 ($x > 0$). The redistribution of the temperature (isothermy) is given by the solution of (5.47):

$$\theta/\theta_1 = 1 - \psi(\mu) \tag{5.48}$$

when T_1 is kept constant ($T \longrightarrow T_1$) and $x > 0$ or

$$\theta/\theta_1 = \tfrac{1}{2}[1 + \psi(\mu)] \tag{5.49}$$

if $T \longrightarrow \frac{1}{2}(T_1 + T_2)$ and $-\infty < x < +\infty$. In (5.48) and (5.49),

$$\psi(\mu) = \frac{2}{\sqrt{2\pi}} \int_0^{\mu} e^{-y^2} dy \tag{5.50}$$

* Such an atmosphere does not represent real conditions and has been chosen as a working model.

and

$$\mu = \frac{x}{2}\left(\frac{n}{A_1 T^{1/2} t}\right) \tag{5.51}$$

where t denotes the time; A_1 is given by (5.18) or (5.19).

By using (5.12), we can write the relation between the temperatures if we adopt a definite value for $\psi(\mu)$. Since $\psi(\mu)=0{\cdot}2$ leads to conditions near to isothermy, such a value is adopted and introduced in (5·48) and (5.49). With (5.48), the solutions are:

If $T_2/T_1=0{\cdot}5$, $T/T_1=0{\cdot}91$, ..., if $T_2/T_1=0{\cdot}9$, $T/T_1=0{\cdot}98$, and with (5.49),

if $T_2/T_1=0{\cdot}5$, $T/T_1=0{\cdot}80$, ..., if $T_2/T_1=0{\cdot}9$, $T/T_1=0{\cdot}95$.

It is, therefore, possible to define a time of conduction related to the concentration (see 5.51) and the distance x. Using the constants from the formulas (5.17), (5.18) and (5.19), we obtain the times of conduction t_{sec} in seconds for atomic hydrogen, atomic oxygen and air

$$t_{sec}(\mathrm{H}) = 7{\cdot}72\times10^{-9} n(\mathrm{H})\, x_{km}^2 T^{-1/2} \tag{5.52}$$

$$t_{sec}(\mathrm{O}) = 4{\cdot}41\times10^{-8} n(\mathrm{O})\, x_{km}^2 T^{-1/2} \tag{5.53}$$

$$t_{sec}(N_2, O_2) = 1{\cdot}46\times10^{-7} n(O_2, N_2)\, x_{km}^2 T^{-1/2} \tag{5.54}$$

For a temperature range of 900°K to 2500°K, i.e. for $\bar{T}=(40\pm10)^2$ °K, the following times of conduction in seconds for distances x expressed in km and concentrations in cm^{-3} are obtained:

$$t(\mathrm{H}) = (2\pm0{\cdot}5)\times10^{-10} x_{km}^2 n(\mathrm{H}) \tag{5.55}$$

$$t(\mathrm{O}) = (1{\cdot}2\pm0{\cdot}3)\times10^{-9} x_{km}^2 n(\mathrm{O}) \tag{5.56}$$

$$t(\mathrm{O_2}, \mathrm{N_2}) = (3{\cdot}8\pm10)\times10^{-9} x_{km}^2 n(\mathrm{O_2}, \mathrm{N_2}) \tag{5.57}$$

If a horizontal discontinuity exists in the horizontal temperature distribution it cannot be maintained at the highest altitudes. For example, the isothermy should be reached after about 4 hr at a distance of 1000 km in an atmosphere (600–700 km) with concentration of 10^7 oxygen atoms cm^{-3}. This means that a difference of temperature in latitude can only be maintained during a very short time at high altitudes. Furthermore, a seasonal effect cannot be related to differences of temperatures at high altitudes. All conditions are related to the diurnal variations which are important according to the results of the analysis made in Section 5.4.

If we consider that horizontal layers of a certain thickness x occur with different temperatures T_1 and T_2, it is possible to consider the time of conduction which ought to be taken into account to reach isothermy. For example, King-Hele and Walker[42] in their atmospheric model have introduced a peak in the scale height at 210 km ($H=47$ km) and a minimum at 250 km ($H=40$ km). For a molecular nitrogen atmosphere, equation (5.47) shows that

$$\theta/\theta_1 = \tfrac{1}{2}\psi(\mu) \tag{5.58}$$

where

$$\mu^2 = a^2 n / A_1 T^{1/2} t \tag{5.59}$$

if a is the thickness of the layer with initial temperature T_1 ($x<a$) and initial temperature T_2 for $-a<x<a$.

Therefore the isothermy is reached ($\mu=0{\cdot}2$ and $900°\mathrm{K}\leqslant\bar{T}\leqslant1600°\mathrm{K}$) when the time of conduction t_{sec} is

$$t_{sec} = 1{\cdot}7\times10^{-8} a_{km}^2 n \tag{5.60}$$

Since the concentration is about 3×10^9 cm^{-3} near 250 km and about four times more near 200 km, the temperature discontinuity will disappear in about 2×10^4 sec between 200 km and 280 km. These results, related to those obtained for the horizontal heat transport, clearly show that discontinuities in the thermospheric temperatures cannot be permanent.

6. THERMOSPHERIC CONDITIONS AND ATMOSPHERIC MODELS

Heat conduction times which are appropriate to the thermosphere are between 12 and 24 hr since the diurnal variation resulting from the heating by electromagnetic radiation is the more pronounced effect. Any gradient of temperature which is not maintained by an external heat flow has a lifetime of less than one night above 350 km, and in a dark atmosphere such an altitude can be considered as being in the neighborhood of the thermopause level. In a sunlit atmosphere its altitude will depend on the solar energy available and the variable temperature of the isothermal region will be a function of

the temperature gradient between 100 km and 200 km. As we have shown previously a change of the gradient below 200 km does not greatly change the atmospheric density at 200 km, but the atmosphere above 300 km is strongly affected. However, a general effect associated to the solar activity (27-day period and sunspot cycle), producing a modification of boundary conditions in the *E*-layer will affect the atmosphere near 200 km with amplified effects at increasing altitudes.

A "corpuscular effect" will also have a detectable lifetime of the order of one day since it will be involved in the effect of the diurnal variation. On the other hand, such an effect will be much easier to detect in a dark atmosphere when the temperature reaches its normal minimum.

The diurnal variation has another important effect on the altitude of the conventional exosphere, because differences of several hundreds of km make it impossible to determine relations connected with the outer atmosphere on a permanent basis. It is, therefore, necessary to take into account diurnal and solar variations modifying simultaneously the values of the density and the temperature. It should be understood that a diurnal variation of diffusion should exist up to the highest altitudes and that the ratio N_2/O must vary between day and night. On the other hand, the upper part of the ionosphere should be subject to very important diurnal and solar effects. Since the recombination in the *F*-layer depends on reactions related to the temperature, important diurnal variations must occur in the rate coefficients. Above the *F2*-peak there must be a diurnal variation in the scale height associated with the vertical distribution of electrons subject to diffusion.

If it is relatively easy to represent atmospheric conditions in a dark atmosphere where the thermopause is at a relatively low altitude, it is very difficult to find an adequate picture of the atmosphere under sunlit conditions until the solar spectrum is known exactly.

An atmospheric model based solely on a heat flow by conduction from outside the atmosphere[37] is an extreme example which cannot be associated with a diurnal variation of the density. Such a heating could be easily explained by the Chapman[8,9] process, suggesting that the solar corona extends to the limits of the terrestrial atmosphere. However, it must be considered as a very small fraction in relation to the heating due to the ultra-violet radiation and, in fact, it is negligible in comparison. In the same way, corpuscular and hydromagnetic effects are not important under normal conditions, but disturbed conditions associated with magnetic storms, must lead to transient atmospheric conditions in which the temperature of the isothermal layer is higher than that in the night-time atmosphere.

If it is clear that the scale height increases with height and that its gradient may reach a peak at a certain altitude, it is also evident that heat conduction does not permit having a temperature gradient increasing with height; even when the temperature increases up to the highest altitudes. For this reason, the atmospheric models such as those of Mikhnevich *et al.*[32], Champion and Minzner[7], Kallmann[36], etc., do not represent real physical conditions.

In the Mikhnevich model the temperature gradient increases with the altitude by $1°K\ km^{-1}$ at 250 km, 2° at 300 km, 3° at 350 km, 4° at 425 km, 6° at 450 km and 7° at 500 km. Such an increase of the temperature gradient would represent a downward heat flow at 500 km at least ten times greater than the heat flow at 250 km. No permanent physical process could explain such a result.

In the Champion and Minzner model, the temperature difference by steps of 100 km from 200 km to 700 km changes successively from 19°K to 57°K, 96°K, 115°K and 121°K. Such an increase of the temperature gradient with height associated with an isothermy between 210 and 260 km would lead also to a downward heat flow; maximum at the highest levels and disappearing near 250 km. Hence the high temperatures near 700 km cannot result from the vertical distribution deduced by Champion and Minzner[7].

In summary, all atmospheric models involving an increase of the gradient of temperature with

height cannot be accepted. Such models lead to the ratio of heat flows E and E_0,

$$\frac{E}{E_0} = \frac{AT^{1/2}dT/dz}{A_0T_0^{1/2}dT_0/dz} \tag{5.61}$$

in which E is for a height z and E_0 is for a height $z=0$. Since $AT^{1/2} > A_0T_0^{1/2}$ and $E \leqslant E_0$, if there is no radiation loss, dT/dz must be less than dT_0/dz. On the other hand, since the conduction of heat is a very rapid process at highest altitudes there is no way to avoid conclusions obtained from (5.61).

Considering now that the tendency of the atmosphere after sunset is to attain isothermy, a night-time model of the atmosphere must be represented at the highest levels by a quasi-isothermal atmosphere in which the constituents are subject to diffusion. In Section 2.1 (see Fig. 2), it has been shown that an isothermal atmosphere subject to diffusion is not far from a quasi-isothermal atmosphere agreeing with night-time satellite data. Furthermore, in Section 3, it has been found that the density (expression 3.4) at 150 km is not sensitive to the temperature gradient when constant boundary conditions are assumed at 120 km. For these two reasons it can be assumed that atmospheric models for night-time conditions at the highest altitudes must be considered in such a way that the temperature is nearly constant and that diffusion affects the atmospheric scale height.

In such circumstances, an important parameter is the ratio of the densities of molecules and atoms. We start from conditions given by (3.4), i.e. at 150 km,

total density: $\rho = 1{\cdot}5 \times 10^{-12}\ \mathrm{g\,cm^{-3}}$ (5·62)

temperature: $725^\circ\mathrm{K} \leqslant T \leqslant 1650^\circ\mathrm{K}$ (5.63)

concentrations: $n(\mathrm{N_2}) = 2{\cdot}6 \times 10^{10}\ \mathrm{cm^{-3}}$ (5.64)

$n(\mathrm{O}) = 3{\cdot}3 \times 10^{9}\ \mathrm{cm^{-3}}$

$n(\mathrm{O_2}) = 5{\cdot}0 \times 10^{9}\ \mathrm{cm^{-3}}$

leading to the approximate ratios

$$n(M) : n(\mathrm{N_2}) : n(\mathrm{O_2}) : n(\mathrm{O}) = 1 : 0{\cdot}76 : 0{\cdot}14 : 0{\cdot}10 \tag{5.65}$$

The effect of diffusion is to lead, at a certain altitude, to the following relation between the densities

$$\rho(\mathrm{O_2}, \mathrm{N_2}) = \rho(\mathrm{O}) \tag{5.66}$$

corresponding to a certain total density ρ. Computation shows that for the conditions (5.65), (5.66) is obtained where the total density has a definite value, i.e.

$$\rho = (4{\cdot}45 \pm 0{\cdot}05) \times 10^{-15}\ \mathrm{g\,cm^{-3}} \tag{5·67}$$

This important relation (5.67) leads to the immediate conclusion that the temperature of the isothermal atmosphere can be found when the altitude of $\rho(\mathrm{O_2}, \mathrm{N_2}) = \rho(\mathrm{O})$ is known.

The following results leading to a "thermo-isobaric relation" have been obtained:

Table 1.

Altitude in km of $\rho(\mathrm{O_2}, \mathrm{N_2}) = \rho(\mathrm{O})$	Temperature (°K)
300	725
350	958
400	1191
450	1424
500	1657

Fig. 14 shows the thermo-isobaric relation, i.e. the relationship between the temperature of the isothermal layer and the altitude of the constant density given by (5.67) corresponding to the same densities for atoms and molecules. The concentrations are approximately as follows:

$$n(\mathrm{N_2}) = 4{\cdot}4 \times 10^{7}\ \mathrm{cm^{-3}};\quad n(\mathrm{O_2}) = 3{\cdot}3 \times 10^{6}\ \mathrm{cm^{-3}};$$
$$n(\mathrm{O}) = 8{\cdot}6 \times 10^{7}\ \mathrm{cm^{-3}} \tag{5.68}$$

The vertical distribution of the scale height is shown in Fig. 15 for various isothermal atmospheres. An increase of H from 50 km at 250 km to 100 km at 1000 km is given for a temperature of the order of 1400°K. Such an increase results only from the decrease of the molecular mass with altitude. It can also be seen that the fluctuations of the scale height deduced by King-Hele and Walker[42] are anomalous variations of the scale heights in an isothermal atmosphere around 1200°K; not too different from the temperature $T = 1250^\circ\mathrm{K}$ deduced (Fig. 14) from the thermo-isobaric relation.

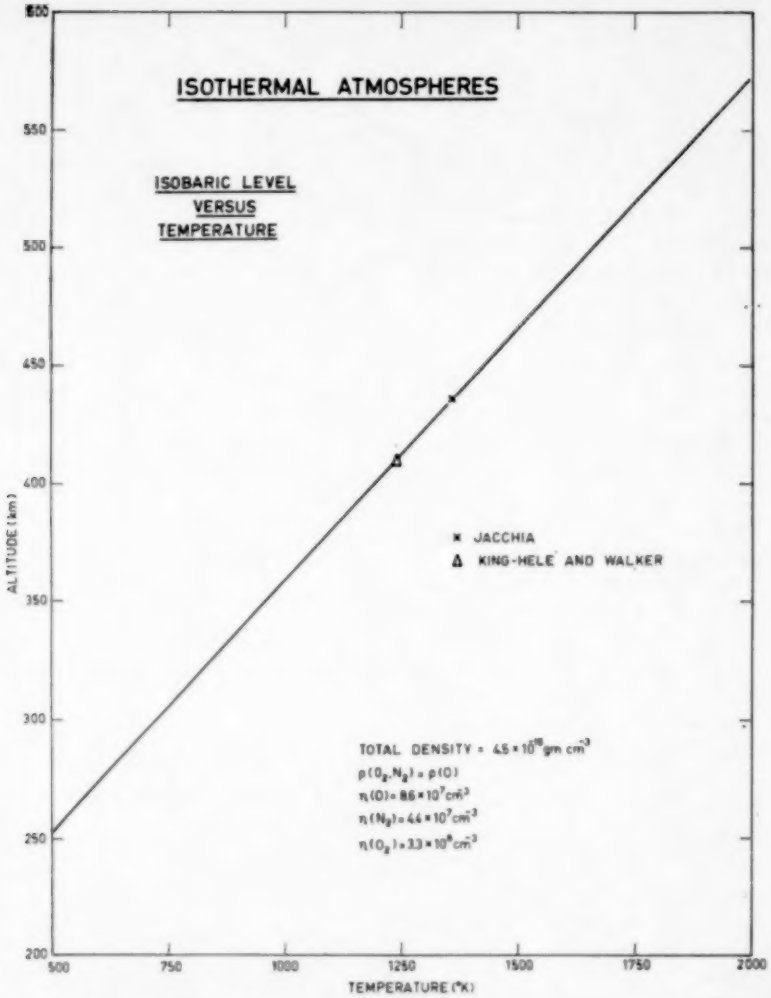


Fig. 14. Temperature–altitude relation for a total density $\rho = 4{\cdot}5 \times 10^{-15}$ gm cm^{-3} where ρ (O_2, N_2) = ρ (O) in an isothermal atmosphere. The knowledge of the altitude of the isobaric level leads to a determination of the temperature and therefore of the vertical distribution of the density.

Fig. 16 shows the vertical distribution of the density for which the constant density defined by (5.67) corresponds to the equality of the density of molecules and atoms. As was shown before (Fig. 2) the curves leading to (5.66) at 400 km and 450 km are not far from the night-time densities deduced by King-Hele and Walker[42] and Jacchia[33], respectively.

It is clear that the preceding determination of night-time conditions of the isothermal layer does not give a complete answer since it depends on boundary conditions for diffusion and temperature. The atmospheric model is consistent in the use of all physical parameters; however, it represents only a guide to the study of atmospheric behavior. In order to show how small

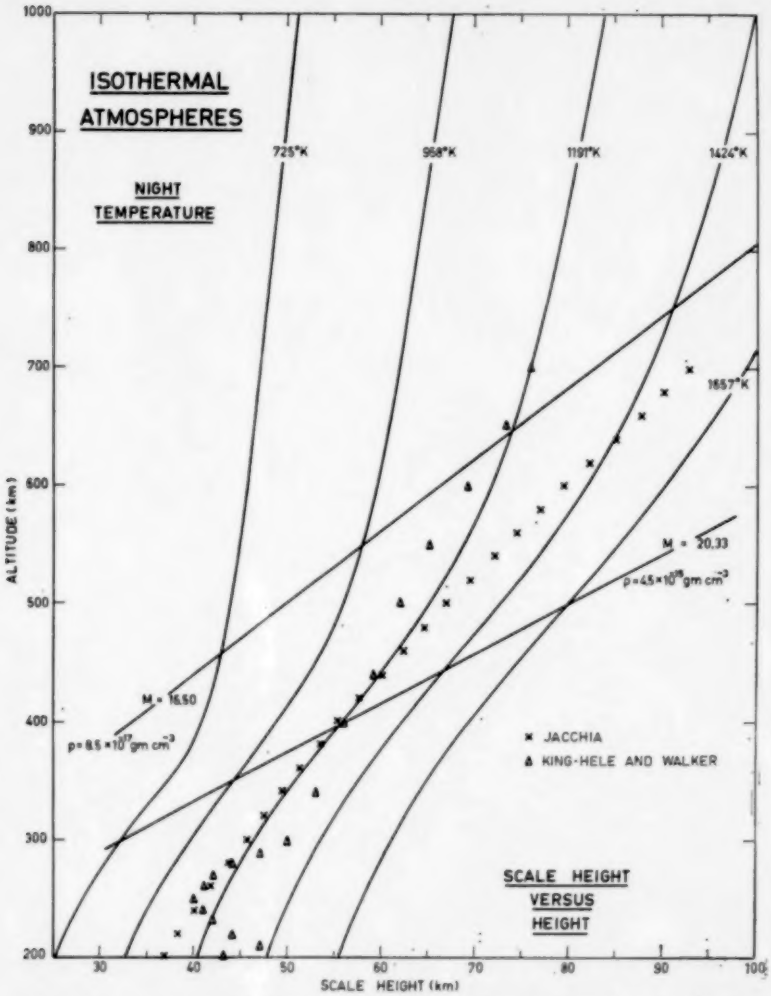


Fig. 15. Scale height–altitude relations for various isothermal atmospheres. Scale height values deduced by Jacchia'[33] and King-Hele and Walker[42] are shown with their different vertical distribution.

differences in physical conditions can modify the conclusions, instead of (5.66), we assume

$$\rho(O_2, O) = \rho(N_2) \qquad (5.69)$$

corresponding again for all temperatures to a constant density. The expression (5·69) leads to a different distribution of the density, even if the concentration of atomic oxygen, $n(O) = 8{\cdot}5 \times 10^7$ cm^{-3}, remains the same at the thermo-isobaric level. The vertical distribution of density for various temperatures are shown in Fig. 16 where it is interesting to compare the results represented by the dotted curves with the data shown by continuous curves.

Computation shows that the condition (5.69) is obtained where the total density has the definite value

$$\rho = (4{\cdot}75 \pm 0{\cdot}25) \times 10^{-15} \text{ g cm}^{-3} \qquad (5.70)$$

This thermo-isobaric relation leads to the results of Table 2.

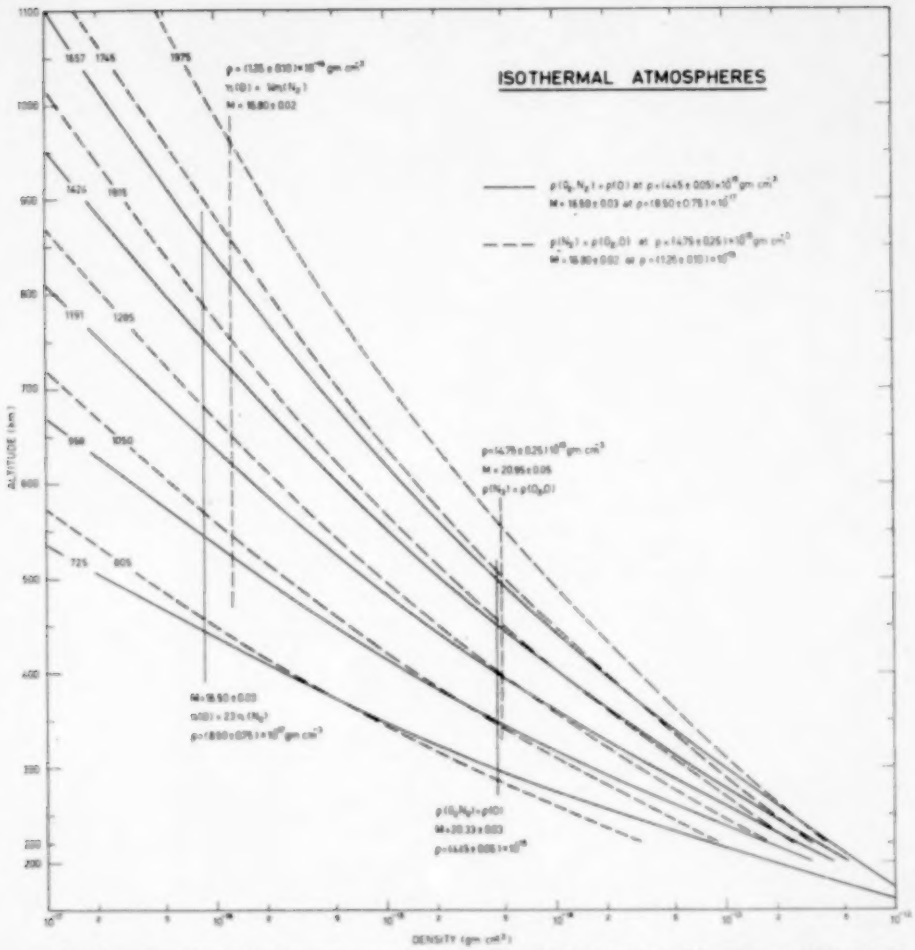


Fig. 16. Densities in isothermal atmospheres. The absolute values of density are determined by conditions ρ (O_2, N_2) = ρ (O) at an isobaric level $\rho = 4{\cdot}5 \times 10^{-15}$ g cm^{-3} and ρ (O_2, O) = ρ (N_2) at an isobaric level $4{\cdot}75 \times 10^{-15}$ g cm^{-3} leading to densities at 220 km corresponding to observed values and to remarkable differences at highest altitudes.

Table 2.

Altitude in km of $\rho(N_2) = \rho(O_2, O)$	Temperature (°K)
350	1050
400	1285
450	1515
500	1745
550	1975

The main conclusions are given in Tables 3(a) and 3(b).

Table 3(a)

Density at 220 km 10^{-13} g cm^{-3}	Isothermal Atmosphere	
	(°K) T_1	(°K) T_2
(1) 0·32	—	805
(2) 0·85 ±0·5	725	1050
(3) 1·7 ±0·1	958	1285
(4) 2·5 ±0·1	1190	1515
(5) 3·2 ±0·1	1424	1745
(6) 3·9 ±0·1	1657	1975

Table 3(b)

	$\rho_1 = 4{\cdot}45\times10^{-15}$ Altitude (km)	$\rho_2 = 4{\cdot}75\times10^{-15}$ Altitude (km)	$\rho_1 = 8{\cdot}5\times10^{-17}$ Altitude (km)	$\rho_2 = 1{\cdot}25\times10^{-16}$ Altitude (km)
(1)		300	—	450
(2)	300	350	450	550
(3)	350	400	550	650
(4)	400	450	650	750
(5)	450	500	750	850
(6)	500	550	850	950
	$M =$	$M =$	$M =$	$M =$
	$20{\cdot}33\pm0{\cdot}03$	$20{\cdot}95\pm0{\cdot}05$	$16{\cdot}50\pm0{\cdot}03$	$16{\cdot}80\pm0{\cdot}02$

The main conclusions are given in Tables 3(a) and 3(b).

Starting from thermo-isobaric levels leading to $4{\cdot}5\times10^{-15}$ g cm^{-3} and $4{\cdot}75\times10^{-15}$ g cm^{-3}, corresponding to $\rho(O_2, N_2)=\rho(O)$ and $\rho(O_2, O)=\rho(N_2)$, respectively, i.e. mean molecular masses $M=20{\cdot}33$ and $M=20{\cdot}95$, the same altitudes are obtained if the temperatures difference is of the order of 100°K. An altitude difference of 50 km corresponds to isothermal atmospheres for which the temperatures differ from each other by about 300°K. In such a case they lead to the same density at 220 km; at the highest altitudes, a difference of more than 100 km is required to reach the same density.

Such large differences show that the structure of the high atmosphere is very sensitive to diffusion and there is a possibility of finding physical conditions which can represent the conditions of a dark atmosphere. From the preceding table and from satellite data, the following data may be taken as a guide (see preceding table): $\rho(220\text{ km})=(2{\cdot}5\pm0{\cdot}7)\times10^{-13}$ for $960°K\leqslant T\leqslant1425°K$ leading to $\rho=4{\cdot}5\times10^{-15}$ g cm^{-3} for 350 km$\leqslant z\leqslant$450 km and $\rho=8{\cdot}5\times10^{-17}$ g cm^{-3} for 550 km$\leqslant z\leqslant$750 km. The altitudes for $\rho=4{\cdot}5\times10^{-15}$ g cm^{-3} and $\rho=8{\cdot}5\times10^{-17}$ g cm^{-3} are related directly to the temperature. Furthermore, the concentrations are given by (5.68) where $\rho=4{\cdot}5\times10^{-15}$ g cm^{-3} and $n(O)=23n(N_2)$ where $\rho=8{\cdot}5\times10^{-15}$ g cm^{-3}. Finally, the mean molecular masses are approximately $M=20{\cdot}33$ at $\rho=4{\cdot}5\times10^{-15}$ g cm^{-3} and $M=16{\cdot}50$ at $\rho=8{\cdot}5\times10^{-17}$ g cm^{-3}.

These various values will serve as a guide in the analysis of night-time conditions in the ionosphere, of airglow and auroras keeping in mind that in a normal dark atmosphere during maximum sunspot conditions the temperature of the isothermal atmosphere may reach values between 1200°K and 1400°K. However, values of the order of 1000°K can be easily accepted.

The problem of the sunlit atmospheres must be studied under various conditions depending on the ultra-violet solar radiation available and on the vertical distribution of its absorption.

Acknowledgments—The research reported in this work has been sponsored by the Geophysical Research Directorate of the Air Force Cambirdge Research Center, Air Research and Development Command, under Contract AF 19 (604)-4563.

The author is grateful to the Astrophysical Observatory of the Smithsonian Institution for supplying the satellite data.

REFERENCES

1. R. Arnoldy, R. Hoffman, and J. R. Winckler, *Proc. COSPAR Symposium, Nice*, p. 877. North Holland Company, Amsterdam (1960).
2. J. Bartels, *Naturwiss.* **45**, 181 (1959).
3. D. R. Bates, *The Earth as a Planet* (Edited by G. P. Kuiper), Chapter 12. Univ. of Chicago Press, Chicago (1954).
4. D. R. Bates, *Proc. Roy. Soc. London*, **253**, 451 (1959).
5. R. E. Briggs, *Smith. Astrophys. Obs. Sp. Rep.* **30**, 9 (1959).
6. E. T. Byram, T. A. Chubb and H. Friedman, *The Threshold of Space* (Edited by M. Zelikoff), p. 211. Pergamon, London (1957).
7. K. W. Champion and R. A. Minzner, *Planet Space Sci.* **1**, 259 (1959).
8. S. Chapman, *Smith. Contr. Astrophys.* **2**, 1 (1957).
9. S. Chapman, *Annales Géophys.* **15**, 434 (1959).
10. S. Chapman and J. Bartels, *Geomagnetism*, p. 897. Oxford Univ. Press (1940).
11. S. Chapman and T. G. Cowling, *The Mathematical Theory of Non Uniform Gases*. Cambridge Univ. Press (1939).
12. K. D. Cole, *Nature* **186**, 874 (1960).
13. A. J. Dessler, *J. Geophys. Res.*, **64**, 397 (1959).
14. A. J. Dessler, *Nature*, **184**, 261 (1959).
15. A. J. Dessler and E. H. Vestine, *J. Geophys. Res.*, **65**, 1069 (1960).
16. P. E. El'Yasberg, *Iskusstvennye Sputniki Zemli*, **1**, 21 (1958).
17. G. V. Groves, *Nature*, **181**, 1055 (1958).
18. G. V. Groves, *Nature*, **182**, 1533 (1958).
19. G. V. Groves, *Nature*, **184**, 179 (1959).
20. G. V. Groves, *Proc. Roy. Soc., London*, **252**, 16 (1959).
21. G. V. Groves, *Proc. Roy. Soc., London*, **252**, 28 (1959).

22. I. Harris and R. Jastrow, *Science*, **127**, 471 (1958).
23. I. Harris and R. Jastrow, *Science*, **128**, 421 (1958).
24. R. Horowitz and H. E. LaGow, *J. Geophys. Res.* **62**, 57 (1957).
25. R. Horowitz, H. E. LaGow and J. F. Guiliani, *J. Geophys. Res.*, **64**, 2287 (1959).
26. L. G. Jacchia, *Smith. Astrophys. Obs. Sp. Rep.*, **12**, 30 (1958).
27. L. G. Jacchia, *Smith. Astrophys. Obs. Sp. Rep.* **13**, 1 (1958).
28. L. G. Jacchia, *Sky and Telescope*, **17**, 278 (1958).
29. L. G. Jacchia, *Smith. Astrophys. Obs., Sp. Rep.*, **20**, 5 (1959).
30. L. G. Jacchia, *Nature*, **183**, 526 (1959).
31. L. G. Jacchia, *Nature*, **183**, 1662 (1959).
32. L. G. Jacchia, *Smith. Astrophys. Obs. Sp. Rep.* **29** (1959).
33. L. G. Jacchia, *Smith. Astrophys. Obs. Sp. Rep.* **39** (1960).
34. L. G. Jacchia and R. E. Briggs, *Smith. Astrophys. Obs. Sp. Rep.* **18**, 9 (1958).
35. R. Jastrow, *J. Geophys. Res.* **64**, 1789 (1959).
36. H. Korf Kallmann, *J. Geophys. Res.* **64**, 615 (1959).
37. D. G. King-Hele, *Nature*, **183**, 1224; **184**, 1267 (1959).
38. D. G. King-Hele, *Progress in the Astronautical Sciences*, to be published (1960).
39. D. G. King-Hele and D. M. C. Walker, *Nature*, **182**, 860 (1958).
40. D. G. King-Hele and D. M. C. Walker, *Nature*, **183**, 527 (1959).
41. D. G. King-Hele and D. M. C. Walker, *Nature*, **185**, 727 (1960).
42. D. G. King-Hele and D. M. C. Walker, *Nature*, **18**, 928 (1960).
43. Y. Kozai, private communication.
44. V. I. Krassovski, *J. Amer. Rocket Soc.* **30**, 27 (1960).
45. J. E. Kupperian, Jr., E. T. Byram, H. Friedman and Unzicker, *CSAGI Meeting, Moscow* (1958).
46. H. E. LaGow, R. Horowitz and J. Ainsworth, *Planet. Space Sci.* **2**, 33 (1959).
47. M. L. Lidov, *Iskusstvennye Sputniki Zemli*, **1**, 9 (1958).
48. A. N. Lowan, *J. Geophys. Res.* **60**, 421 (1955).
49. B. R. May and D. E. Smith, *Nature*, **184**, 765 (1959).
50. L. R. Davis, O. E. Berg and L. H. Meredith, *Proc. COSPAR Symposium, Nice*, p. 721. North Holland, Amsterdam (1960).
51. V. V. Mikhnevich, *Iskusstvennye Sputniki Zemli*, **2**, 26 (1958).
52. V. V. Mikhnevich, B. S. Danilin, A. I. Repnev and V. A. Sokolov, *Iskusstvennye Sputniki Zemli*, **3**, 84 (1959).
53. Mullard Radio Astronomy Observatory, *Nature*, **180**, 879 (1957).
54. E. P. Ney, J. R. Winckler and P. S. Freier, *Phys. Rev. Letters*, **3**, 183 (1959).
55. M. Nicolet, *Science*, **127**, 1317 (1958).
56. M. Nicolet, *Aeronomic Chemical Reactions, in Physics and Medicine of the Atmosphere and Space* (Edited by O. O. Benson, Jr. and H. Strughold), Chap. 2. John Wiley, New York (1960).
57. M. Nicolet, *Ann. Géophys.* **15**, 1 (1959).
58. M. Nicolet, *J. Geophys Res.* **64**, 2092 (1959).
59. M. Nicolet, *Upper Atmosphere* (Edited by Ratcliffe), Chapter II, Academic Press, New York (1960).
60. M. Nicolet, *Proc. COSPAR Symposium, Nice*, p. 46. North Holland, Amsterdam (1960).
61. H. K. Paetzold, *Naturwissenschaften*, **13**, 416 (1958).
62. H. K. Paetzold, *Planet. Space Sci.* **1**, 115 (1959).
63. H. K. Paetzold, *Raketentechnik und Raumfahrtforschung*, **2**, 45 (1959).
64. H. K. Paetzold, *Zeit. angew. Phys.* **11**, 234 (1959).
65. H. K. Paetzold, *J. Atm. Terr. Phys.* **16**, 259 (1959).
66. H. K. Paetzold and H. Zschörner, *Proc. COSPAR Symposium, Nice*, p. 24. North Holland Company, Amsterdam (1960).
67. W. Priester, *Naturwiss Meschaften*, **46**, 197 (1959).
68. W. Priester, H. G. Bennewitz and P. Lengruesser, *Westdeutscher Verlag*, (1958).
69. W. Priester and H. A. Martin, *Mitteilungen der Universitäts—Sternwarte Bonn*, No. 29 (1959).
70. Royal Aircraft Establishment, *Nature*, **180**, 937 (1957).
71. G. F. Schilling and T. E. Sterne, *J. Geophys. Res.* **64**, 1 (1959).
72. G. F. Schilling and C. A. Whitney, *Smith. Astrophys. Obs. Sp. Rep.* **18**, 13 (1958).
73. G. F. Schilling and C. A. Whitney, *Planet. Space Sci.* **1**, 136 (1959).
74. G. F. Schilling, C. A. Whitney and B. M. Folkart, *Smith. Astrophys. Obs. Sp. Rep.* **14**, 32 (1958).
75. L. I. Sedov, *Iskusstvennye Sputniki Zemli*, **2**, 1 (1958).
76. I. S. Shklovsky and V. G. Kurt, *Iskusstvennye Sputniki Zemli*, **3**, 66 (1959).
77. J. W. Siry, *Planet. Space Sci.* **1**, 184 (1959).
78. T. E. Sterne, *Science*, **128**, 420 (1958).
79. T. E. Sterne, *Phys. Fluids*, **1**, 165 (1958).
80. T. E. Sterne, *J. Amer. Rocket Soc.* **4**, 777 (1959).
81. T. E. Sterne and G. F. Schilling, *Smith. Contr. Astrophys.* **2**, 207 (1958).
82. T. E. Sterne, B. M. Folkart and G. F. Schilling, *Smith. Contr. Astrophys.* **2**, 275 (1958).
83. J. Van Allen, *J. Geophys. Res.*, **64**, 1683 (1959).
84. J. W. Warwick, *Planet. Space Sci.* **1**, 43 (1959).
85. J. R. Winckler, *Non-Relativistic Protons from Solar Flares* (*Symposium on Upper Atmosphere Solar Relations, Chicago, Dec.* 1959), to be published.
86. S. P. Wyatt, *Nature*, **184**, 351 (1959).
87. R. A. Minzner, K. S. W. Champion and H. L. Pond, *Air Force Surveys in Geophysics* No. 115 (1959).

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 33-45. Printed in Great Britain.

STUDIES OF AURORAL ECHOES–I

L. HARANG and J. TRÖIM

Norwegian Defence Research Establishment, Division of Telecommunication, Kjeller, Lilleström, Norway

(*Received* 10 *December* 1959)

Abstract—The drift of the auroral ionization in an E–W direction is studied by means of auroral echoes on a frequency of 40 Mc/s. Two identical channels, with aerials at a certain divergence symmetrical to north, are used. The crossing over of the blobs, giving receding and approaching ranges on the records is studied, and the E–W drift is measured. Two series of observations have been made, at Kjeller (60°N) and at Tromsö (70°N). The mean diurnal curves of E–W motions at these two places are given. The diurnal variation at Kjeller is similar to the drifts recorded at Jodrell Bank. At the auroral station Tromsö the drifts are more irregular. There is a prevailing westerly drift between 17·00 hr–24·00 hr MET and an easterly drift between 24·00 hr–05·00 hr MET. Examples of drift effects using interferometers are shown.

The angle of arrival of the auroral echoes has been studied at Kjeller and Tromsö by an interference method. The aerial, used both for transmission and reception, slides up and down a mast, and the interference fringes between the direct and the reflected ray is measured. The angle of arrival, lying in the interval 6°–20°, shows the same dependence of range at Kjeller and Tromsö. The height of the auroral ionization is determined and the aspect sensitivity for reflection is studied. The angle of arrival of E_s–propagated ground scatter is measured and is shown to have about twice the value of the angle of arrival of the auroral echoes.

1. DRIFT MOTIONS OF AURORAL IONIZATION

Introduction

Studies of auroral echoes in frequency ranges of 40–100 Mc/s have indicated the presence of horizontal drifts of the scattering areas of auroral ionization. Different methods have been used for analysing these drift motions. The simplest method would be to use a pencil-beam aerial directed at low elevation sweeping continuously in azimuth, and to display the auroral echoes on a PPI-screen. Bullough and Kaiser[1] have successfully used this method and shown that the pattern from auroral ionization, observed at a middle latitude station like Jodrell Bank, agrees with the hypothesis that the ionization is distributed along an arc which follows the parallels of geomagnetic latitude. In some cases it has also been possible on the PPI-screen to follow successive displacements in azimuth of a scattering area. Another method which turned out to be more convenient for continuous observation was introduced by Bullough and Kaiser[2]. A narrow-beam aerial is directed towards NW or NE, if the drift of the auroral ionization crossing the aerial pattern follows the geomagnetic lines of latitude, the variation in range with time will indicate either a receding or approaching motion of the auroral ionization from which the E–W drift can be calculated. A third method is to analyse the frequency spectrum of the CW-auroral reflections received by aerials directed towards NW or NE, from which the Doppler-displacement can be calculated (McNamara and Currie[3], Nichols[4]). All methods indicate a drift effect; the motions are directed either towards east or west depending on the time of the day, with speeds ranging from 0·5 to 3 km/sec. It has further been pointed out by Kaiser[2] that there seems to be a well-defined systematic variation in the direction of motion, which changes from a mean of about 0·6 km/sec westwards during the evening maximum to about 0·6 km/sec eastwards during the morning. Kaiser *et al.*[5, 6] have also given detailed analysis of the geomagnetic variations during individual days on

which aurorae appeared, and has shown that for positive and negative geomagnetic bays the drift of the auroral ionization was towards west and east respectively.

Diurnal variation of the auroral echoes

Two series of measurements were made: the first series at Kjeller ($\varphi = 59° 58'$ N, $\lambda = 11° 16'$ E) and the other at Tromsö ($\varphi = 69° 40'$ N, $\lambda = 18° 57'$ E). Kjeller lies 1300–1500 km to the south of the maximum auroral zone. The auroral echoes were observed towards the north at ranges between 400–900 km and the echoes appearing at Kjeller were therefore from the southern edge of the auroral zone. At Tromsö the distance to the "normal" auroral zone is about 300 km, and with the ranges usually appearing in a northerly direction the echoes recorded would be from the inner edge of the auroral zone.

It was found that the echoes appearing at Kjeller were of a far more regular character than at Tromsö and special attention will therefore be paid to the Kjeller recordings.

The period of observations at Kjeller covered the 8 winter months October 1957–May 1958. At Tromsö recordings were made on 35 days only during the summer months July to August 1958. The instrumental equipment at Kjeller and Tromsö was identical. In Fig. 1 is shown the number of hours during which auroral echoes have been recorded. As a comparison, the diurnal curve for Jodrell Bank according to Bullough and Kaiser[2] is given.

The diurnal variations show that the auroral echoes are predominantly a night phenomenon at middle latitude stations like Kjeller and Jodrell Bank. The diurnal variation at the two places is parallel and the two maxima appear at the same local time. In the diurnal curves no attention has been paid to the strength of the echoes, and the higher frequency of appearance at Kjeller will tend to smooth out the secondary night minimum. The rate of appearance of the echoes at Kjeller is, according to Fig. 1, about thirty times the rate of appearance at Jodrell Bank. This figure must, however, be considered with caution. The Jodrell Bank observations were made during a period with low solar

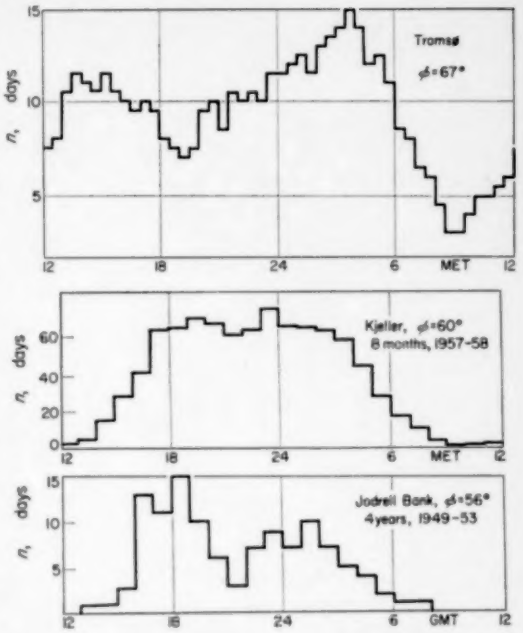


Fig. 1. Diurnal appearance of auroral echoes on 40 Mc/s at Tromsö and Kjeller. n is the number of days during which echoes have been recorded. The diurnal curve from Jodrell Bank (Bullough and Kaiser[2]) is given as comparison. The geomagnetic latitudes, for the three stations are respectively 67°, 60° and 56°.

activity during which the auroral activity is low at middle latitude stations, whereas the Kjeller observations were made during a year of maximum solar activity. The parameters of the instrumentation at the two stations seem, however, to have been of the same magnitude.

At Kjeller auroral echoes appeared almost on every night. At Tromsö the echoes appeared most frequently during the night, but there is also a high frequency of occurrence during the day which does not occur at middle latitude stations. There is also a shift in the time of occurrence of the maxima and the minima. At Tromsö there is a high frequency of occurrence between 1200 hr–1600 hr MET, whereas at Kjeller and Jodrell Bank the frequency during these hours is very low. It must, however, be remembered that the echoes recorded at Tromsö were reflected from

regions lying 200–400 km north of the "normal" auroral zone and, further, the records were all made in summer during midnight-sun conditions.

Equipment for drift studies

The two-channel equipment used worked on 40·2 Mc/s and 39·6 Mc/s respectively. Fig. 2 shows the principle of the equipment. At Kjeller the two Yagi aerials were directed to 12° E and 40° W of the N–S direction respectively. According to Störmer[7] this should be approximately symmetric to the direction of auroral

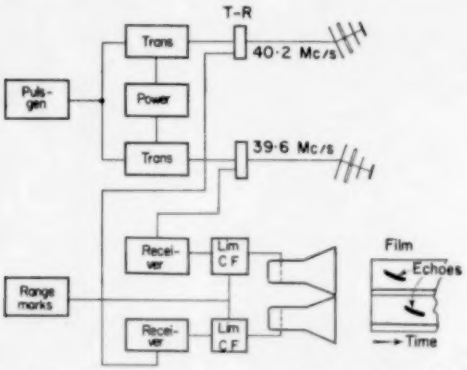


Fig. 2. The two-channel equipment for drift recordings.

arcs appearing over Southern Norway. At Tromsö the experiments were started with a pair of Yagi aerials, but due to the irregular character of the echoes within the auroral zone it was necessary to use aerials with narrower beam-widths and also to decrease the angular separation between the aerials.

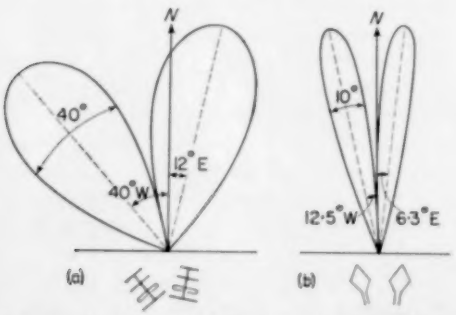


Fig. 3. Aerials for drift studies (a) Kjeller: two 3-element Yagi aerials. (b) Tromsö: two rhombic aerials.

An indication of the variation in drift directions can be obtained from an inspection of the two-channel records. On Fig. 4 are indicated the various characters of range variations on the records which are to be expected from the east and west aerials, depending on the drift directions, and records from Tromsö

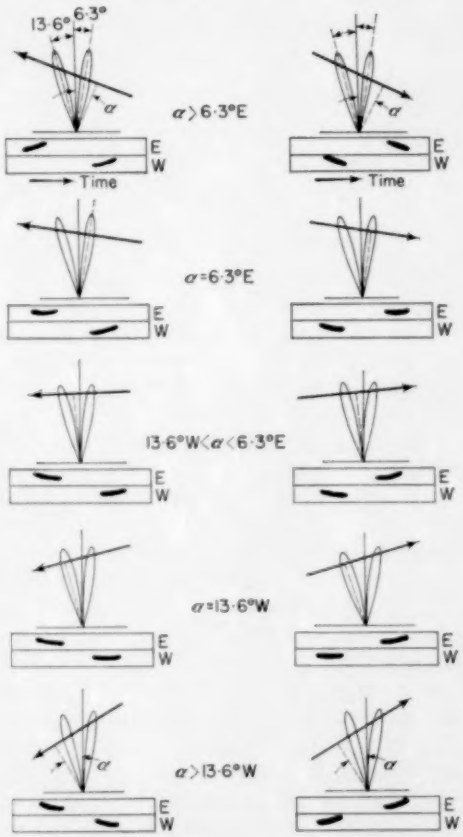


Fig. 4. Different types of range variations recorded on the two channels, depending on the direction of drift. From the type of range variations the direction of the drift can be estimated.

show examples of all these types of range variations. At Tromsö some few cases were also observed where the ionization must have drifted through only one single lobe, indicating an almost N–S drift.

The beam-width in the horizontal and vertical planes are of decisive importance with regard to the character and range of the echoes recorded. At Kjeller horizontal Yagi aerials were used, placed at an height of 1λ above the ground. At Tromsö the height of the rhombic aerials was also 1λ above the ground. The antenna patterns in the horizontal plane could easily be measured. In the vertical plane where the effect of ground reflections is of decisive importance, the position of the main lobe was controlled by trapezium, and from the horizontal sides, knowing the speed of drift and the range, the extension in E–W direction of the blob can be estimated. Only a few cases appeared in which the records showed the simple character of individual patches as in Fig. 6. In most cases the records show a continuous echo pattern with receding or approaching fine structure.

Recordings of drift at Kjeller

At Kjeller the two Yagi aerials were pointed

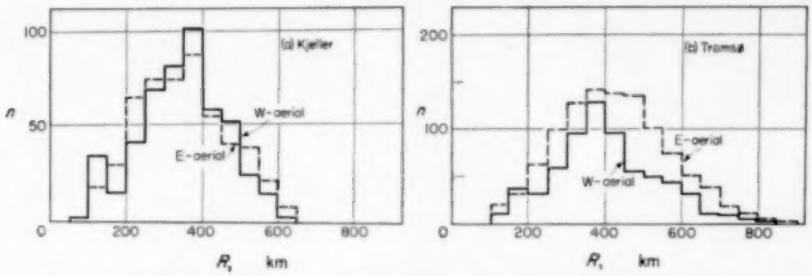


Fig. 5. Number of meteor echoes counted (n) versus range (R). (a) Yagi aerials at Kjeller, (b) rhombic aerials at Tromso.

counting meteor echoes. Fig. 5(a, b) shows the number of meteors as a function of range recorded through the Yagi and rhombic aerials.

Assuming a mean height of reflection of the meteors of 100 km, the main lobe of the Yagi aerials will cut the ionized region at a range of about 350 km, and the rhombic aerials at about 400 km. The two Yagi aerials seem to have identical vertical antenna lobes, whereas for the two rhombic aerials the eastern has a broader vertical antenna lobe than the western, but the maxima of the lobes go through the same region.

The different types of range variations indicated on Fig. 4 can be discussed from the actual records and finer details of the drift effects can be demonstrated. If a blob elongated in E–W direction is crossing the lobe of the aerial, it is easy to show that the range variations on the records will, in an ideal case, take the form of a trapezium. Fig. 6(a, b) shows the range variations to be expected and recorded. It is evident that the time of passage through the aerial pattern of the front of the blob is determined by the left and right sides of the

in a northerly direction with an angular separation of 52°, and with the line of symmetry pointing against 15° W. On Figs. 7 and 8 are shown a number of typical records. The drift directions can be estimated from the approaching and receding radial velocities observed on both channels. In some cases it is possible to identify the same scattering area entering the two antenna lobes in succession.

If one assumes that the drift of the scattering areas is along the geomagnetic lines of latitude it should be possible, through an analysis of the radial velocities (Kaiser *et al.*[11]), to assess the speed. The echo patterns obtained at Kjeller seem, however, to be considerably more complicated than those obtained at Jodrell Bank—this station lies 4° of geomagnetic latitude south of Kjeller and a quantitative estimate of the change in radial velocities versus time at Kjeller was usually difficult. Moreover, it was often observed during midnight that receding and approaching radial echo structures appeared at the same time. In the analysis of the records we have therefore not tried to estimate the rate

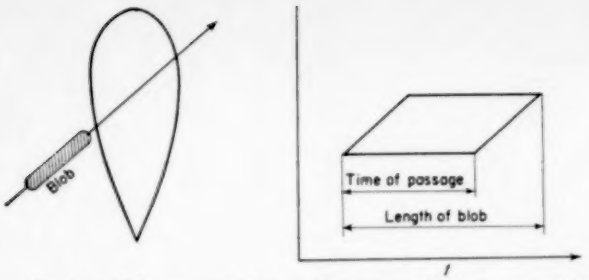
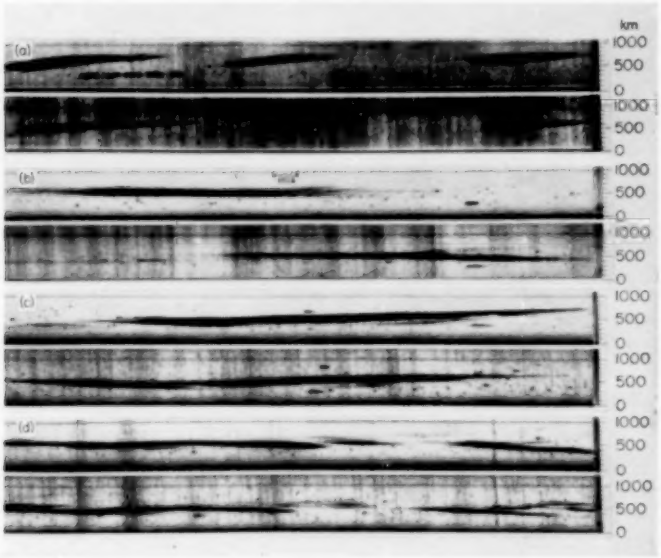


Fig. 6. (a) Range variations to be expected when a single E–W elongated blob crosses the aerial pattern, (b) record of drift of a single blob.



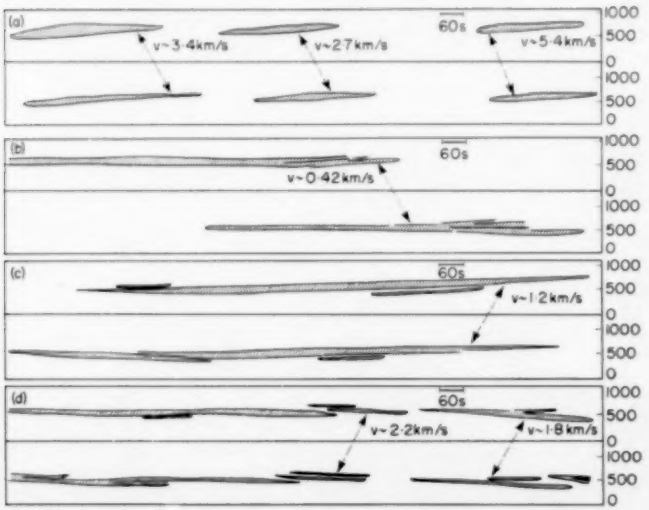


Fig. 7. Drift records from Tromsö of auroral echoes of simple character.

Upper channel: rhombic aerial pointing 6·3°E.
Lower channel: rhombic aerial pointing 12·5°W.

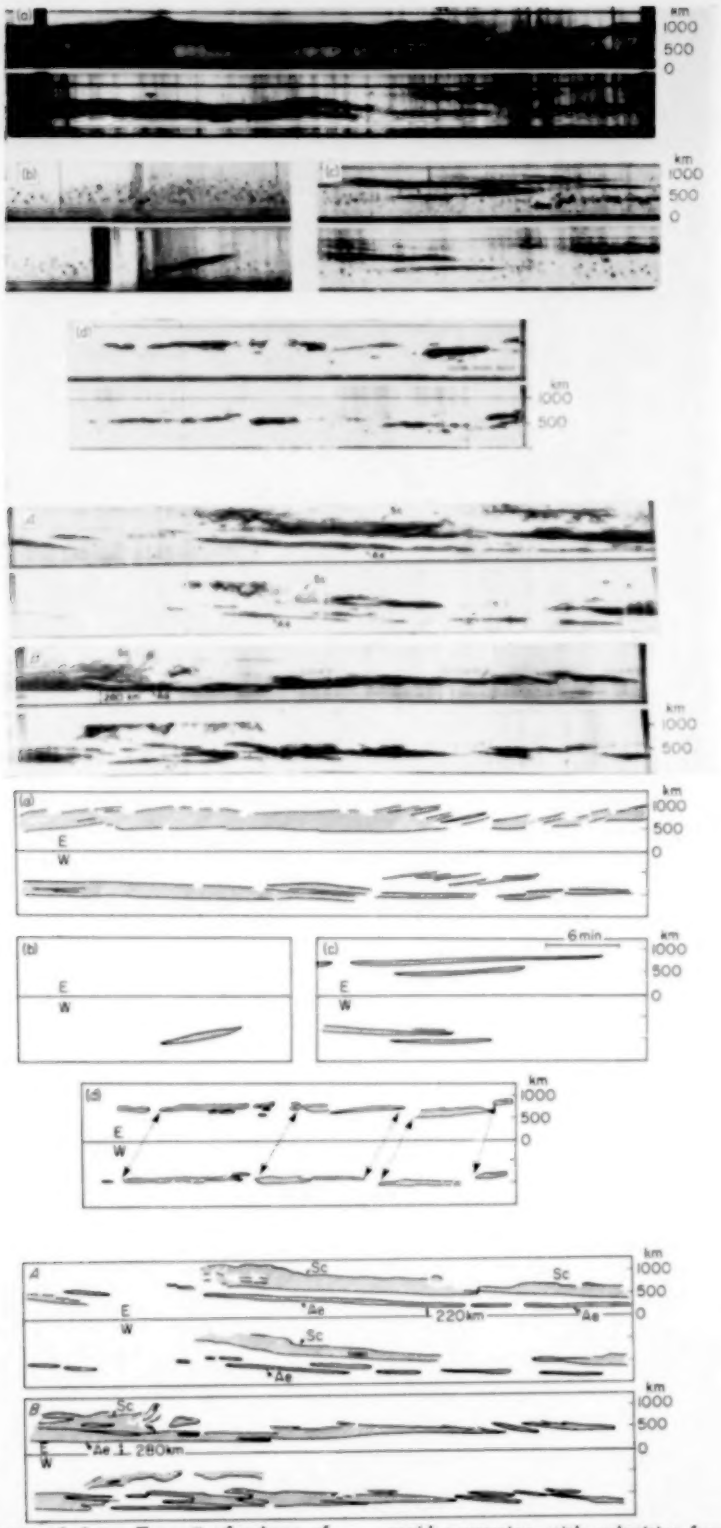


Fig. 8. Drift records from Tromsö of echoes of complex character. (a) Record indicating a multitude of complex movements. (b) Echo appearing on only one channel. Echo (W) showing high radial velocity indicating a large N—S movement. (c) Week echoes moving E with velocities from 1·4 to 2 km/sec. (d) Six blobs appearing in rapid succession with velocities from 1·4 to 4·4 km/sec. (A) and (B) 2-hour record of simultaneous auroral echoes (Ae) and what is believed to be E_s–propagated ground scatter (Sc). A N–S drift of both types is apparent. There is indication of a close connection between the two echo types.

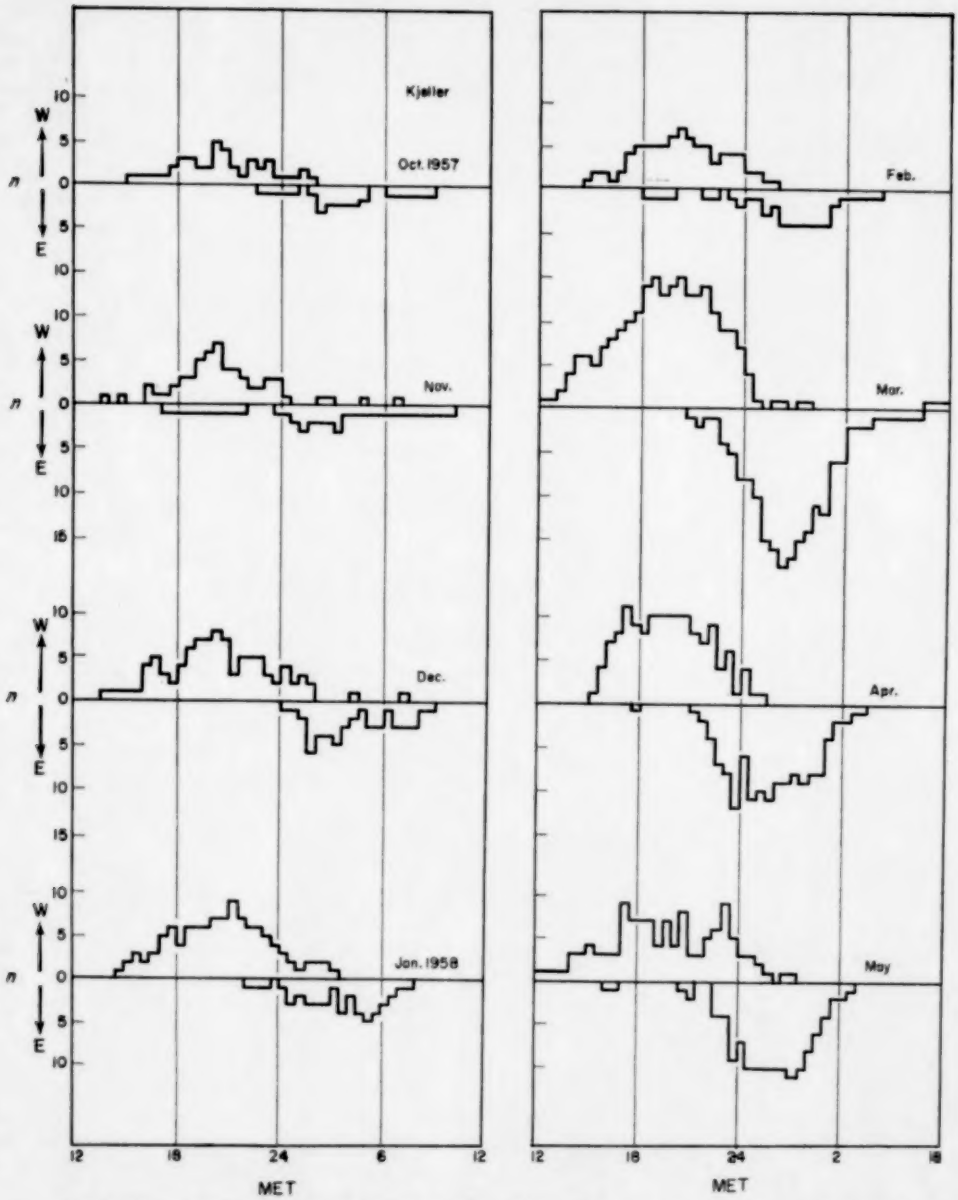


Fig. 9. Mean monthly drift directions at Kjeller of auroral echoes on 40 Mc/s. *n* is the number of days when echoes occurred.

of change in radial velocities, but only differentiated between receding and approaching radial movements, from which the rate of occurrence of westward and eastward motion could be determined. Fig. 9 shows the mean diurnal directions of motion for each month, and in Fig. 10 the mean values of direction for the whole period of 8 months are summed up.

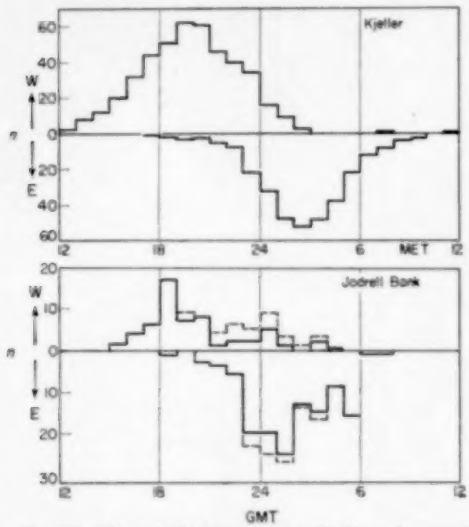


Fig. 10 Mean drift directions at Kjeller of auroral echoes on 40 Mc/s for an 8 month period. The mean drift directions at Jodrell Bank are given according to Bullough and Kaiser[2].

It is evident that between 18.00–22.00 hr MET there is a westerly drift and that between 02.00–06.00 hr MET the drift is eastwards. About midnight there are many cases of eastward and westward drifts, but in the mean there is zero drift. As a comparison the number of westerly and easterly drift directions recorded at Jodrell Bank (see Bullough and Kaiser[2] Fig. 11) are shown in the lower part of the figure, due to the greater number of cases at Kjeller the diurnal curves run smoother than at Jodrell Bank. There is a distinct difference in time for the change over from westward to eastward drift. At Jodrell Bank the change takes place at about 21.30 hr GMT, whereas at Kjeller a similar change takes place at about 24.00 hr MET.

Observations with interferometer

An important question in the theory of blob formation is the size of the blobs. From the double-beam records one gets the impression that auroral ionization is formed, more or less continuously, as cylinders lying in E–W direction and extending over several hundreds of kilometers. In order to obtain a higher resolution of the aerials in the horizontal plane, two receiving Yagi aerials were coupled as an interferometer. Three different base-lengths, 2·4λ, 3·76λ and 7·0λ, were tried. In Fig. 11 three records taken with these spacings of the interferometer are shown.

In the first case (*a*) the records clearly indicate a blob crossing the four lobes on the interferometer. The size of the blob must in this case be of the order of 100 km in an E–W direction. On the records (*b*) and (*c*) there are series of small patches which are interpreted as the crossing over of small blobs imbedded in the great cylinder of ionization. This broken appearance of auroral ionization appeared only on a small number of records. A further development of this interferometer technique should be promising, not only for study of blob size but also for velocity determination. It may be possible that a closer study will reveal different velocities and individual movements of microstructure within the ionization.

Recordings of drift at Tromsö

Preliminary tests at Tromsö showed that the drift effects at this station which lies close to the auroral zone were considerably more irregular than at Kjeller, and the Yagi aerials had to be replaced by narrow-beam rhombics. The diurnal variation of the echoes, based on 35 days continuous recordings, are shown in Fig. 1. The direction of drift is shown in Fig. 12, and as a comparison the drift measurements which have been made at College (Alaska) by Nichols[4] using the Doppler shift method are shown. Tromsö and College both lie close to the auroral zone and should represent similar geophysical conditions with respect to the current systems in the ionosphere. The drifts at these polar stations have not the regularity of character observed at the middle

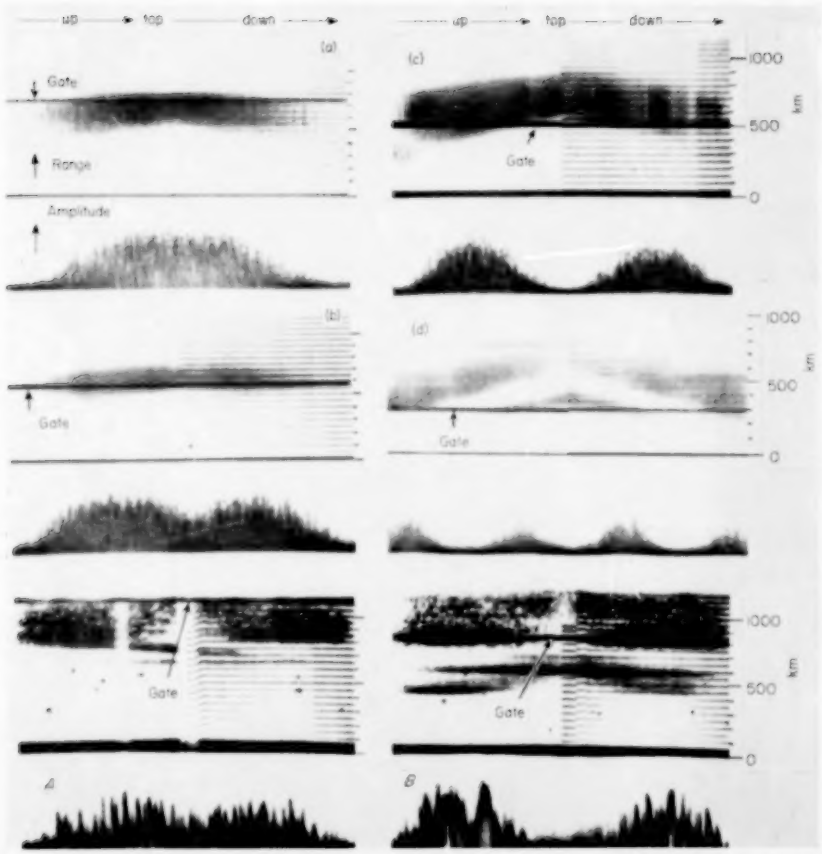


Fig. 18. Records for determination of angle of arrival θ of auroral echoes taken at Tromsö. θ is determined from the position of the amplitude maximum:

(a) gate at $R = 720$ km, measured $\theta = 6°$
(b) gate at $R = 575$ km, measured $\theta = 7{\cdot}5°$
(c) gate at $R = 490$ km, measured $\theta = 9°$
(d) gate at $R = 290$ km, measured $\theta = 20°$

The two lower records, (A) and (B), are examples of E_s–propagated ground-scatter, which gave an angle of arrival of about twice the value of the auroral echoes.

(A) Gate at $R = 1150$ km, measured $\theta = 7{\cdot}5°$
(B) Gate at $R = 850$ km, measured $\theta = 11$.

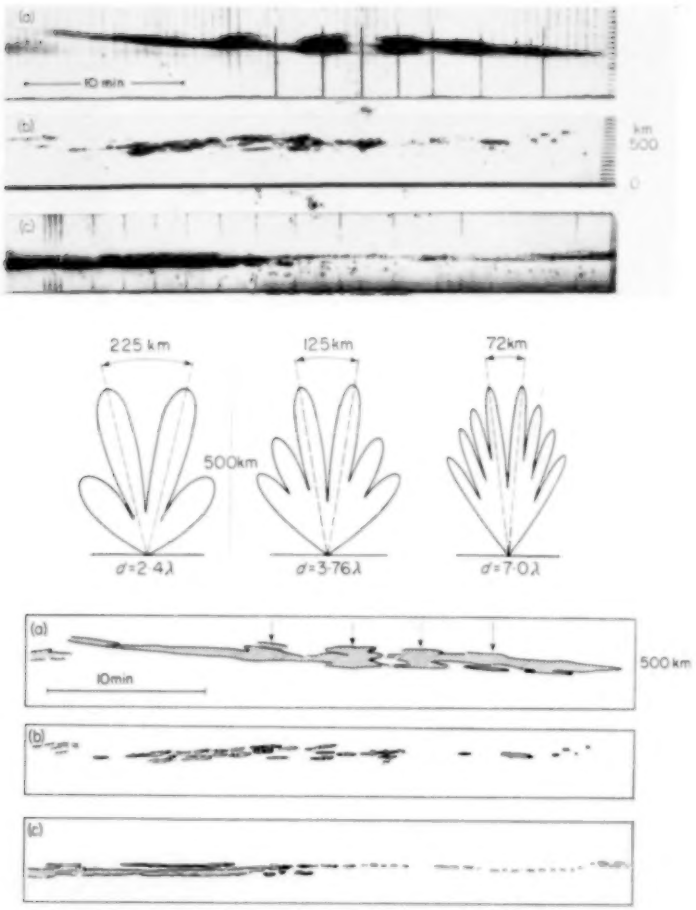


Fig. 11. Records of blob crossing the lobes of the interferometer. Three different spacings have been used: d=2·4λ, 3·76λ and 7·0λ.

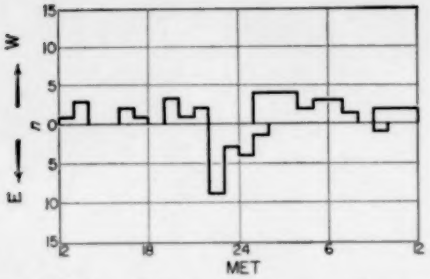


Fig. 12. Mean E–W drift directions at Tromsö of auroral echoes on 40 Mc/s. n is the number of days. Mean drift at College (Alaska) on 41 Mc/s (after Nichols[4]).

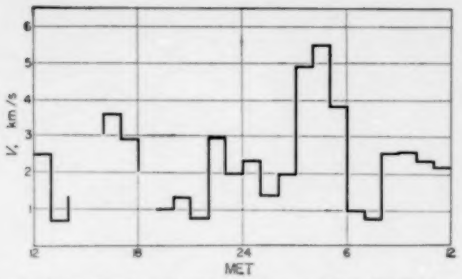


Fig. 14. Drift velocities of auroral ionization at Tromsö.

latitude stations Jodrell Bank and Kjeller. It is apparent, however, that there is an easterly motion at 22.00–24.00 hr at both places.

Directions and speed of the blobs at Tromsö

From the discussion in connection with Fig. 4 it is evident that it is possible in many cases to obtain an estimate of the azimuth of the direction of drift by inspecting the character of the echo pattern on the double-beam records. In Fig. 13 are shown the drift directions and speeds from Tromsö records. There is a considerable variation in the drift directions at Tromsö, and on some records echoes appeared only on one beam, indicating that in this case the drift must have been almost in a N–S direction.

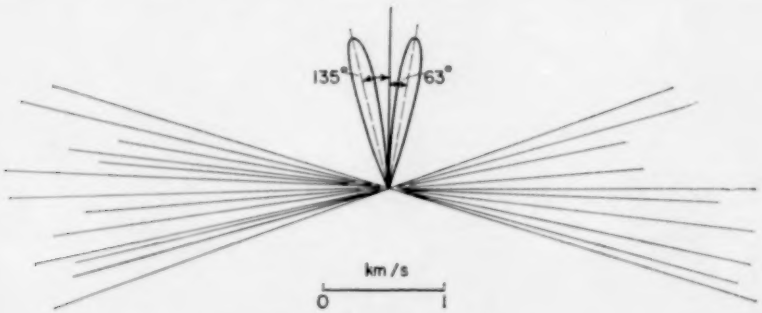


Fig. 13. Drift directions at Tromsö.

In a number of cases the drift velocities have been calculated from the transit time of crossing over the lobes. In Fig. 14 are given the results from a series of measurements at Tromsö. The drift velocities range from 0·6 to 4–5 km/sec. This is a considerably greater value than observed by Nichols[4] at College using the Doppler method, where drift velocities in the range 0·4–1·5 km/sec were observed.

Direction of drift, geomagnetic activity and drift of auroral forms

Bullough *et al.*[6] have pointed out that the drift towards the west in the afternoon and the reversal towards the east around midnight coincides with the changes in direction of the currents which produce the geomagnetic storms. For a station situated to the south of the auroral zone the reversal of the currents are indicated through the reversal of the horizontal and vertical components of the geomagnetic perturbing vector. Bullough *et al.* used the vertical component as indicator for reversal. A similar study has been made by Lyon and Kavadas[8] using the horizontal component of the perturbing vector as comparison. It was found that

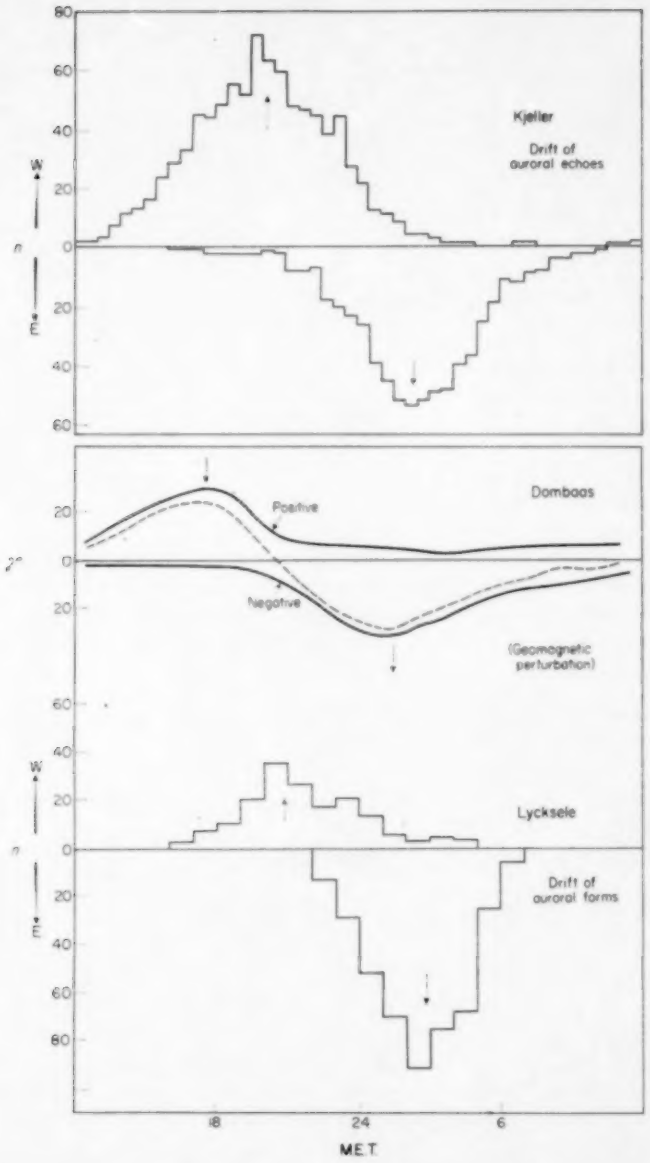


Fig. 15. Mean diurnal variation of (a) E–W drift of ionized areas determined from echoes on 40 Mc/s at Kjeller, (b) the horizontal component of the geomagnetic perturbing vector at Dombaas and Tromsö, (c) E–W drift of auroral forms determined from all-sky photos at Lycksele (after Stoffregen).

the drifts towards the west in the afternoon nearly coincided with the positive value of the perturbing vector, and the drift towards the east after midnight with the negative value of the horizontal perturbing vector.

Geomagnetic recordings were made at the two stations Dombaas and Tromsö; Dombaas lies about 250 km and Tromsö about 1100 km to the north of Kjeller. The auroral echoes with ranges 400–900 km will thus lie between these two geomagnetic stations.

Recordings of aurorae with all-sky cameras in Scandinavia during the IGY were made at the stations Tromsö, Kiruna and Lycksele. The auroral echoes recorded at Kjeller appearing with normal ranges would be reflected from ionized areas lying in zenith at the region of Lycksele, and it is thus of special interest to compare the drifts from the echo records at Kjeller with the E–W drifts of auroral forms observed from all-sky photos taken at Lycksele.

In Fig. 15 is shown the diurnal variation of the E–W drift direction observed at Kjeller, the mean diurnal variation of the positive and negative values of the horizontal perturbing vector observed at Dombaas and Tromsö, and the mean diurnal variation of E–W drifts of auroral forms at Lycksele ($\varphi = 64°$ N) determined from all-sky camera* records.

It is evident that the frequency of drifts towards the east during and after midnight coincides exactly with the variations of the negative (decreasing) values of the geomagnetic perturbing vector. In the frequency curve of motions towards the west in the afternoon there seems, however, to be a difference in phase—the maximum in the positive disturbance appears 1–2 hr before the maximum in the drift motions. It is of interest to notice that on the curves given by Lyon and Kavadas the same time lag seems to appear.

A comparison between the diurnal variation of the E–W drift directions of auroral ionization determined from echo records and the E–W drift of auroral forms determined from all-sky photos, shows a close similarity. The general trend in the variations is the same and the maxima of occurrence of drifts towards the west and the east appear at the same hours. There thus seems to be a close parallel between the mean diurnal movements in E–W directions of the auroral forms and the ionized areas.

A survey of auroral motions based on all-sky photos taken under and to the south of the auroral zone in northern Canada has been published by Kim and Currie[9] and a diurnal motion of the auroral forms in E–W direction similar to that found at Lycksele seems to appear.

2. ANGLE OF ARRIVAL OF AURORAL ECHOES

Experimental arrangement

The auroral echoes are normally observed in the north with the aerial pointing at an angle of low elevation. Qualitative estimates indicate that the angle of arrival must be less than 15°–20°. In a previous paper[10] a method has been described according to which the angle of arrival has been determined as a function of the echo range. The method has been applied on echoes recorded at Kjeller, but has now been further developed and a series of angle arrival determinations have now also been made at Tromsö. Fig. 16 illustrates the method and Fig. 17 shows the position of the aerial with a small lake being used as reflector.

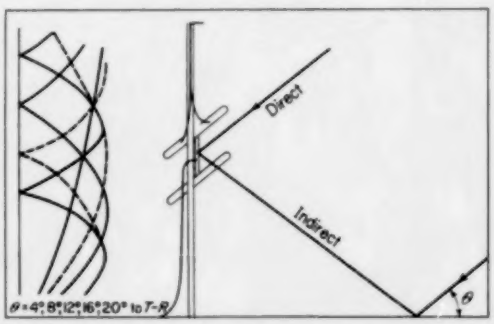


Fig. 16. Method for measuring the angle of arrival of auroral echoes. The transmitting+ receiving aerial slides uniformly up and down. The amplitude variations for different angles of arrival, θ, are indicated to the left.

* All-sky cameras (of the Stoffregen type) have been in operation during the IGY at the Scandinavian stations Tromsö, Kiruna and Lycksele. The records have been worked up by Stoffregen who has kindly supplied the Lycksele-curve of auroral drifts for comparison.

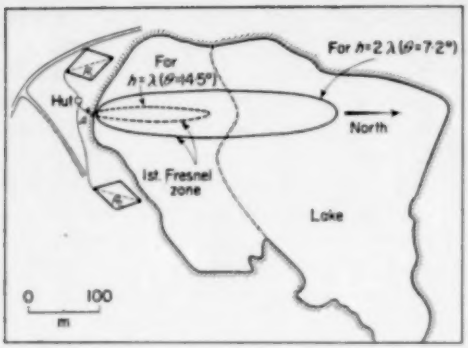


Fig. 17. Position of aerials at Tromsö. The aerial for angle of arrival measurements, A, is placed at the southern edge of a small lake. The positions of the first Fresnel zones are indicated for two values of θ. The two rhombic aerials, R_1 and R_2, were used for drift measurements.

The method is based upon measuring the amplitude of the signal in an aerial moving up and down. With horizontal polarization and complete reflection of the reflected ray the amplitude will vary with height according to the formula $e=2E\sin(2h\sin\theta/\lambda)$, where E is the amplitude of the incoming rays, direct and reflected, h is the height and θ is the angle of arrival. The amplitude variations for different values of θ are indicated to the left on Fig. 16. From the spacings between the maxima (which corresponds to the fringes in Lloyd's fixed mirror experiment in optics) the angle of arrival, θ, can be evaluated.

Due to the considerable width of the auroral echo pattern, a well-defined range for amplitude recording must be selected. This was done by means of a narrow gate set at a selected range. Fig. 18 shows a number of records taken at Tromsö. Two cathode-ray tubes were used, the upper for recording the echo pattern and the position of the gate, the lower for recording the echo amplitude at the range selected through the gate.

The different character of fading of the E_s-propagated ground-scatter and the auroral echoes is apparent.

Observations at Kjeller and Tromsö

A series of measurements of θ at Kjeller and Tromsö are given in Fig. 19; there seems to be the same dependence of θ with R at both places. The ranges, R, varied from 300 km to 850 km and the values of θ varied from 20° to 6° depending on the range.

Three curves have been calculated showing the variation of θ with R if the scattering region is assumed to lie at 70, 100 and 130 km height respectively. For middle ranges, 300–600 km, the height of the scattering region seems to be

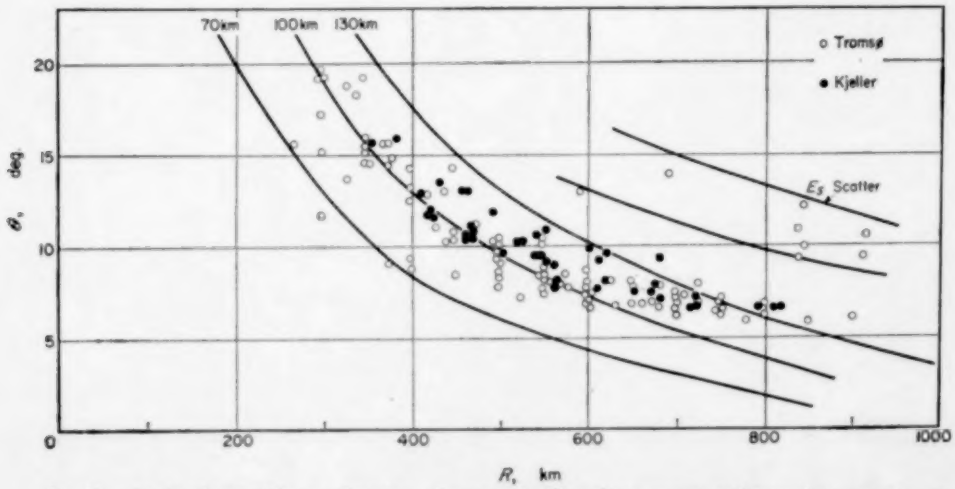


Fig. 19. Angle of arrival θ, as a function of range, R. The three curves indicate the variation of θ with R to be expected if the reflection took place at 70, 100 or 130 km height respectively.

in the region of 100 km. We notice that the decrease of θ with increasing R is slightly less for great ranges, 600–800 km, than for the short and middle ranges. The value of θ for great ranges seems to be somewhat too high.

Such an effect could be expected if there was an appreciable effect of refraction of the waves in the lower atmosphere. The effect of optical refraction is, however, too small to account for this, since for an angle of 6° the effect of refraction will be of the order of 0·2°[11]. We are inclined to assume that the effect may be due partly to inaccuracies of the method at these very low angles of elevation and partly to a strictly fulfilled one should, at Kjeller, record scattering areas from 100 km height at a range of 380 km. For Tromsö the curve for normal incidence on the geomagnetic field lines does not reach the 100 km height for any value of R (see curve A–B on Fig. 20). An analysis of the aspect sensitivity versus range is given in Fig. 21 for Tromsö. At middle ranges, 500–600 km, the condition for perpendicularity is almost fulfilled, the difference being only 2°–3° from normal incidence. At greater ranges and, especially, at smaller ranges down to 300 km there could be a deviation from normal incidence of 6°–9°.

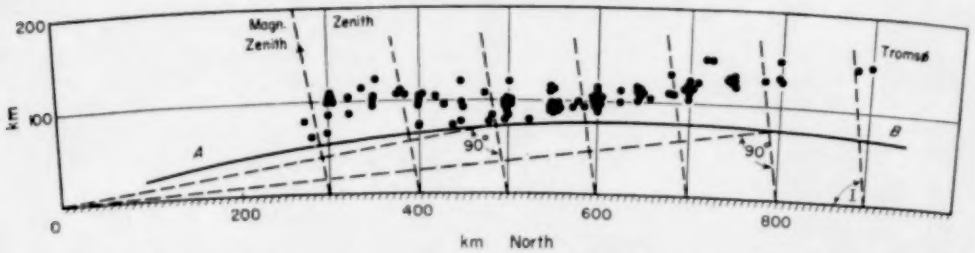


Fig. 20. Apparent height of the scattering region. Using the values of θ and R, the positions of the scattering areas have been plotted in a vertical N–S plane. Observations from Tromsö on 40 Mc/s. The curve A–B indicates the height of reflection to be expected if reflection only took place perpendicular to the geomagnetic field lines.

bending of the ray in the E_s-region formed during the auroral display.

In Fig. 20 the apparent heights of the scattering areas are demonstrated, based on observations at Tromsö. Using the values of θ and R the positions of the reflecting areas have been plotted in a vertical plane lying in N–S direction.

The fact that the auroral echoes are almost exclusively observed at low angles in a northerly direction, led to the assumption of aspect sensitivity of the scattering sources. The model usually accepted is that the scattering areas are made up of ionized columns aligned along the geomagnetic field lines. The scattering process should thus represent an analogy to the scattering of VHF-waves from underdense meteor trails, and one should expect a strong aspect sensitivity since the reflections are strong when the beam is perpendicular to the magnetic field in the scattering area (Chapman[12]): If this assumption of aspect sensitivity should be

The bending of the ray during the passage through the E_s-layer

During auroral activity there will be an appreciable electron density in the E_s-layer which will go down even to 80 km. When penetrating this E_s-layer the 40 Mc/s pulse will therefore be deviated slightly towards the ground. If we knew the vertical electron profile it would be possible to trace the orbit of the pulse for a given angle of arrival. Numerical calculations of the orbits of low angle reflections for various models of electron density profiles have been performed by several authors (Leadabrand[13], Alpert *et al.*[14] and others). Of special interest is the analysis made by Leadabrand of the possible orbits of auroral scatter from and above the F-layer. In this case he analyses the influence of the F-layer on the bending of the ray.

During auroral echo conditions it is assumed that the lower part of the E-layer, even down to

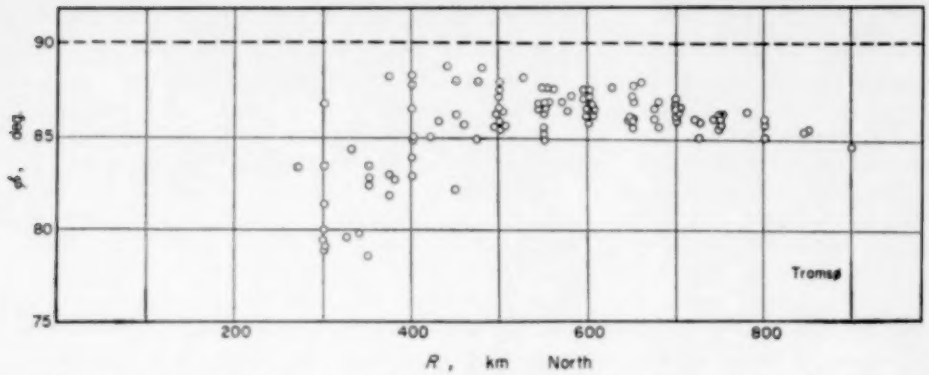


Fig. 21. Aspect sensitivity of auroral echoes at Tromsö. Reflections at normal incidence appear for $\theta = 90°$.

80 km, has a considerable electron density. The shape of the electron density profile towards greater heights is less known, but judging from the luminosity distribution of auroral rays and draperies, one must assume that the ionization will extend upwards several tens of km. Meos and Olving[15] have expressed the view that this ionization may have a considerable effect on the orbit of the ray.

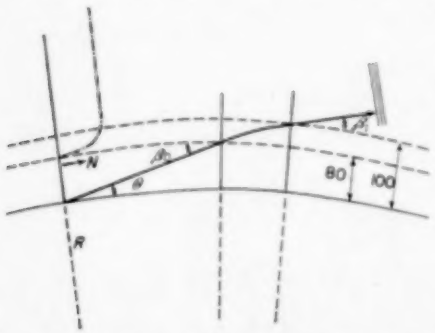


Fig. 22. Bending of a 40 Mc/s ray in an E_s-layer with parabolic electron distribution at the lower edge and constant electron density upwards from electron density maximum.

In order to get numerical estimates of the bending of the ray in an E_s-layer, the following simple model will be considered.

The E_s-layer goes down to 80 km with a parabolic electron density distribution at the lower edge. From electron density maximum, which is assumed to be at 100 km height, the electron density is constant. The ray when entering the E_s-layer, will be bent downwards, but after having passed the electron density maximum the ray will proceed with unchanged direction in an ionosphere of constant electron density. In this case we can use Snell's law for a curved ionosphere (Bouger's law) and we will have

$$\cos\beta_1/\cos\beta_2 = \mu_0 r_0/\mu_1 r_1$$

where

$$r_0 = R + 80\text{ km} = 6447\text{ km}$$

and

$$r_1 = R + 100\text{ km} = 6467\text{ km}.$$

$\mu_0 = 1$ and $\mu_1 = \sqrt{[1-(f/f_c)^2]}$, where f is the frequency of the wave and f_c is the critical frequency of the E_s-layer. For auroral echoes the angle of arrival will lie in the interval 6°–18°, and assuming the following values of $f/f_c = 3$, 5 and 7 (corresponding to $f_c = 13{\cdot}3$, 8·0 and 5·7 Mc/s) the values in the following table are obtained.

It is evident that the depression of the ray may attain a value up to 10° at low angles and for a reasonable value of the maximum electron density in the E_s-layer. The gradual increase in apparent height with range demonstrated in Fig. 22 is therefore explained as being due to the bending of the ray in the E_s-layer. It is also apparent that this bending is sufficient for explaining the deviations from perpendicularity during the scattering process demonstrated on Fig. 23.

		For f_c = 13·3 Mc/s	For f_c = 8·0 Mc/s	For f_c = 5·7 Mc/s	Depression of the ray $\Delta = \beta_0 - \beta_1$		
θ	β_0	β_1	β_1	β_1			
6°	10·8°	—	2·1°	8·4°	—	8·7°	2·4°
10	13·4	—	8·3	11·6	—	5·1	1·8
14	16·6	—	12·8	15·2	—	3·8	1·4
18	20·1	6·7	17·1	18·9	13·4	3·0	1·2

E_s-propagated ground-scatter

In addition to the auroral echoes, a second type of scattered reflections appeared in a number of cases and this must be interpretated as ground-scatter propagated back via E_s-ionization. On some occasions the angle of arrival of these echoes could be determined. The range of this scatter, being of the order 700–1100 km, was considerably greater than for auroral echoes, and there was also a difference in the fading characteristic between the ground-scatter and the auroral echoes. The records *A* and *B* on Fig. 20 illustrate the difference in the fading characteristic between the ground-scatter and the auroral echoes. The ground-scatter had a slow general fading upon which series of rapid fadings are superimposed, whereas the auroral echoes had a very rapid, almost noisy, character of fading. On Fig. 21 are shown some determinations of angle of arrival of the ground-scatter. The observations indicate that for a range of 1000–1100 km there was an angle of arrival of about 10°, which indicates that the ground-scatter must have been reflected back via an E_s-layer at about 100 km height.

The diurnal variation of the appearance of this E_s-propagated ground-scatter received from the north at a range of 800–1000 km, was entirely different from the diurnal variation of the auroral echoes. Fig. 23 shows the diurnal variation; the scatter appeared between 18.00 hr and 24.00 hr MET exclusively.

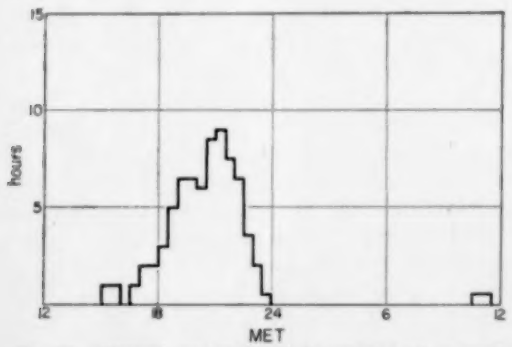


Fig. 23. Diurnal variation of E_s-propagated ground-scatter on 40 Mc/s, received from the north.

Acknowledgment—The authors wish to express their thanks to Mr. R. Brattensborg for his assistance during the observations, and to the Director of the Auroral Observatory at Tromsö, Mr. E. Tönsberg, for hospitality during their stay for observations.

REFERENCES

1. K. Bullough and T. R. Kaiser, *J. Atmos. Terr. Phys.* **5**, 189 (1954). *J. Geophys. Res.* **62**, 297 (1959).
2. K. Bullough and T. R. Kaiser, *J. Atmos. Terr. Phys.* **6**, 198 (1955).
3. A. G. McNamara and B. W. Currie *J. Geophys. Res.* **59**, 279 (1954). A. G. McNamara, *J. Geophys. Res.* **60**, 257 (1955).
4. B. Nichols, *J. Atmos. Terr. Phys.* **11**, 292 (1957); see also *Rep. Geophys. Inst. Univ. Alaska* (1957).
5. T. R. Kaiser, *Ann. Geophys.* **14**, 76 (1958).
6. K. Bullough, T. W. Davidson, T. R. Kaiser and C. D. Watkin, *J. Atmos. Terr. Phys.* **11**, 237 (1957).
7. C. Störmer, *The Polar Aurora*. 101 Clarendon Press, Oxford (1955).
8. G. F. Lyon and A. Kavadas, *Can. J. Phys.* **36**, 1661 (1958).
9. J. S. Kim and B. W. Currie, *Can. J. Phys.* **36**, 160 (1958).
10. L. Harang and J. Troim, *J. Atmos. Terr. Phys.* **14**, 107 (1959).
11. J. L. Pawsey and R. N. Bracewell *Radio Astronomy* p. 342. Clarendon Press, Oxford (1955).
12. S. Chapman, *J. Atmos. Terr. Phys.* **3**, 1 (1952).
13. R. L. Leadabrand, *Techn. Rep. Radio. Propag. Lab., Stanford Univ.* No. 98 (1955).
14. L. Alpert, F. F. Dobiakova, Z. A. Malina and Y. P. Nesterova, *Results of Research in the outer Region of the Ionosphere*. Academy of Sciences, Moscow (1957).
15. J. Meos and S. Olving, *Internal Rep. Res. Lab. Electronics, Chalmers Univ. Technol. Gothenburg* No. 7 (1957).

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 46-49. Printed in Great Britain.

RECENT RESULTS IN THE INVESTIGATION OF THE RELATION BETWEEN LIGHTNING DISCHARGES AND WHISTLERS

HARALD NORINDER and EDGAR KNUDSEN

Institute of High Tension Research, Uppsala University, Sweden

(*Received* 20 *April* 1960)

Abstract—A station for research into the relation between lightning discharges and whistlers was operated near Uppsala and preliminary results obtained were published in this Journal. Further analysis of data obtained has been carried out. It has been possible by harmonic analysis to show that lightning discharges producing whistlers have a pronounced energy peak around 5 kc/s.

From a comparison of time intervals of multiple lightning discharges recorded by a cathode-ray oscillograph and from corresponding records of multiple whistlers on sonagrams full agreement in time variation between the two phenomena was demonstrated.

In a previous article in this Journal[1] the relations between lightning discharges and whistlers were investigated and discussed. Since the article was issued, the authors have had at their disposal two special instruments: a harmonic analyser and a sound-spectrograph. In that way it has been possible to extend the analysis of the observations and the results obtained and reported in the following will in two interesting respects offer a complete account of the relations between lightning discharges and whistlers. In the first respect the frequency distribution of lightning discharges, whether followed by whistlers or not, will be considered. In the second respect, the time relation between multiple lightning discharges and multiple whistlers will be discussed.

In the previous article thunderstorm centres which emitted whistlers were analysed. The mode of procedure was that at the whistler station a cathode-ray oscillographic direction-finder was used to localize thunderstorm centres from which the whistlers were emitted. By an open antenna system at the station in combination with an amplifier and a recording cathode-ray oscillograph the field-force variations from the thunderstorm centres were obtained either from lightning discharges which produced whistlers or from those which lacked whistlers. It happened, for instance, in the beginning of one intense whistler situation that 30 per cent of all oscillographic recordings gave subsequent whistlers; during a later period atmospherics with subsequent whistlers reached about 85 per cent.

In the whistler situations observed the field-force values were measured for the atmospherics recorded, whether followed by whistlers or not, and in the previous article examples were given from which there appeared a very clear tendency as regards the relation of the field-force values to whistlers. It was found that in the region of lowest field-force values atmospherics were not accompanied by whistlers. In an intermediate region on the scale atmospherics occur with or without whistlers. In the higher region of the scale all atmospherics are accompanied by whistlers. Three characteristic groups of field-force variations in atmospherics were consequently obtained: (1) oscillograms of atmospherics not followed by whistlers, (2) oscillograms of atmospherics followed or not followed by whistlers, (3) oscillograms of atmospherics all followed by whistlers. A typical example of the group distribution is given in Fig. 1 in which the black circles mark field-force values followed by whistlers. Systematic comparisons between lightning discharges productive of whistlers and lightning discharges non-productive of whistlers must obviously be made in group (2), where both variation forms occur. In our previous

Fig. 1. Electric field of atmospherics related to whistlers.

article this method was followed and oscillograms of atmospherics from group (2) in Fig. 1 were reproduced which were either followed by whistlers or not followed by them. As already mentioned, it was not possible to present in our article any definite results of a frequency analysis of the respective oscillographic variation types. Preparatory tests showed energy peaks around 5 kc/s, and it may easily be seen that the oscillographic records in the article taken from group (2) in Fig. 1 should first be tested by a harmonic analyser. The results are given in Figs. 2–4, from which there follows a preponderance of frequencies around 5 kc/s for oscillograms of atmospherics followed by whistlers as compared with adjacent oscillograms of atmospherics not followed by whistlers. The same prevalence of frequencies around 5 kc/s has also been obtained by harmonic analysis of other curves belonging to the above mentioned group (3).

In discussions of the relations between multiple lightning discharges and multiple whistlers the possibilities in these relations have very often been presented in a hypothetical way and never before by direct measurements. Obviously the conditions for development of multiple whistlers related to multiple lightning discharges in the same path must be based on time analysis between the two phenomena.

In the investigations presented in the previous article two cathode-ray oscillographs were used for determination of the electric field-force variations from lightning discharges related to whistlers. By a special method[2] one of the

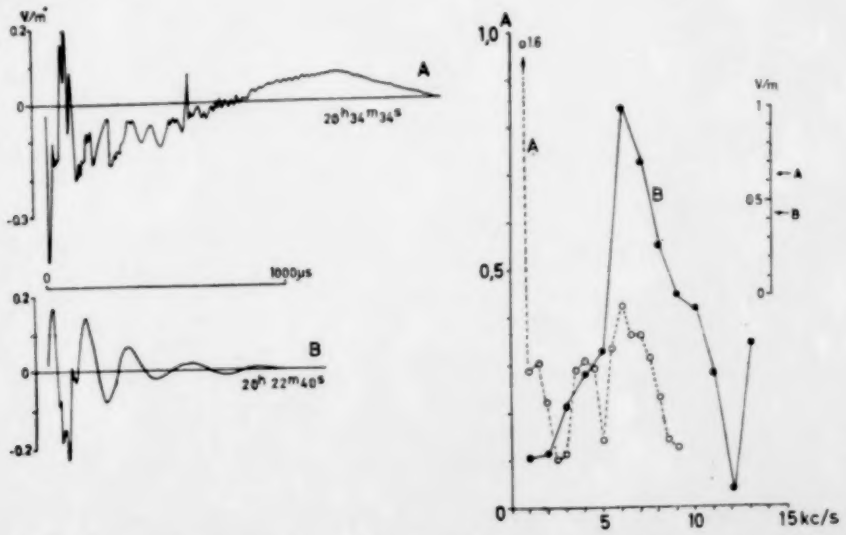


Fig. 2. Harmonic analysis of lightning discharges not followed by whistlers (A), followed by whistlers (B).

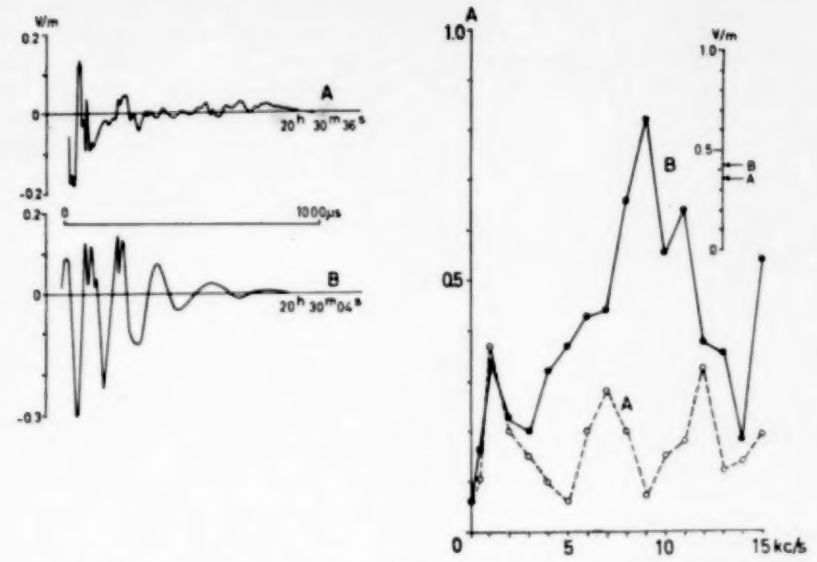


Fig. 3. Harmonic analysis of lightning discharges not followed by whistlers (A), followed by whistlers (B).

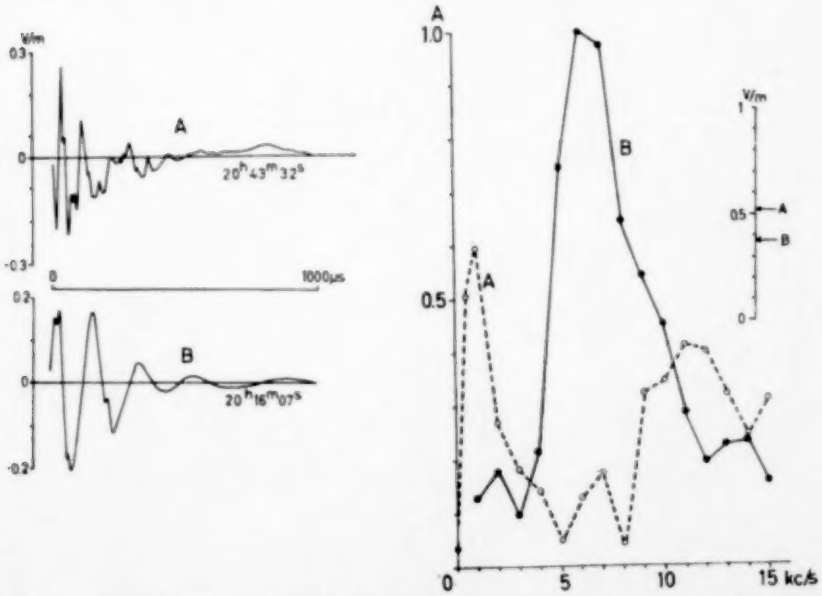


Fig. 4. Harmonic analysis of lightning discharges not followed by whistlers (A), followed by whistlers (B).

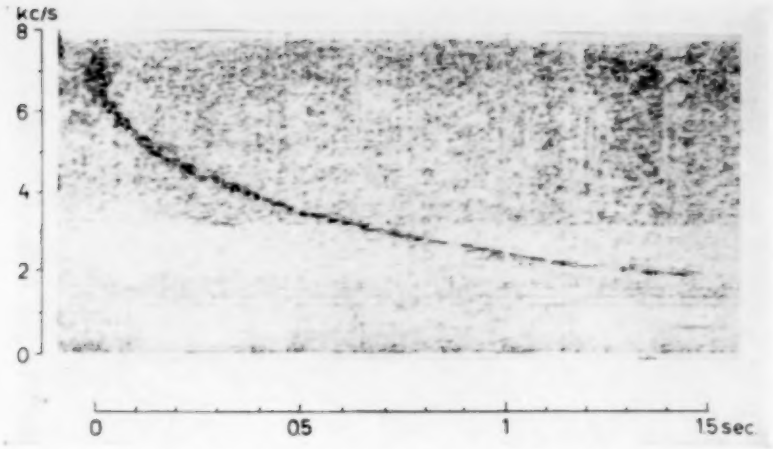


Fig. 5. Sonagram of a single whistler.

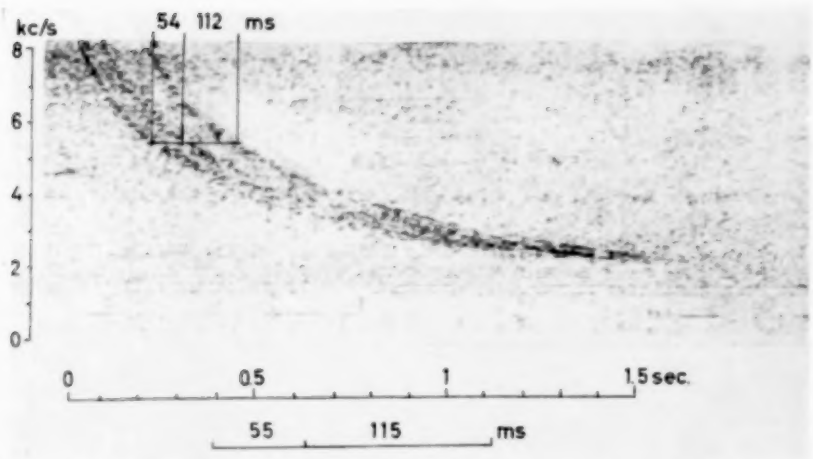


Fig. 6. Sonagram of three multiple whistlers in the same lightning path.

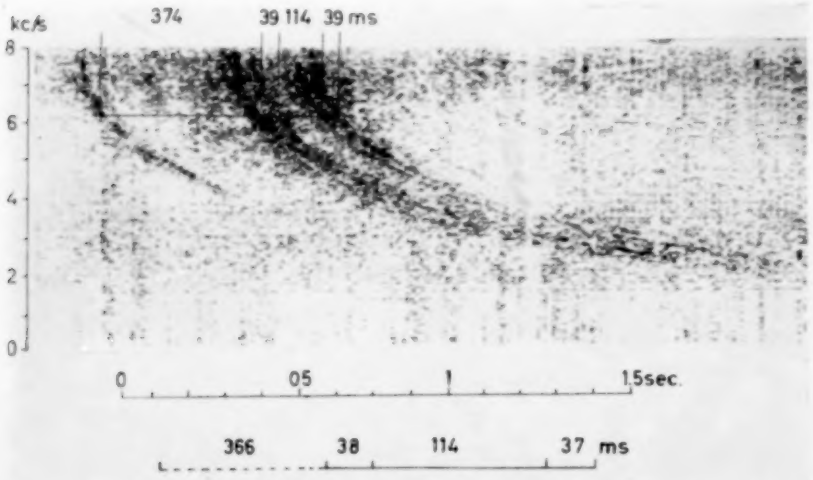


Fig. 7. Sonagram of five multiple whistlers, one followed by long time interval.

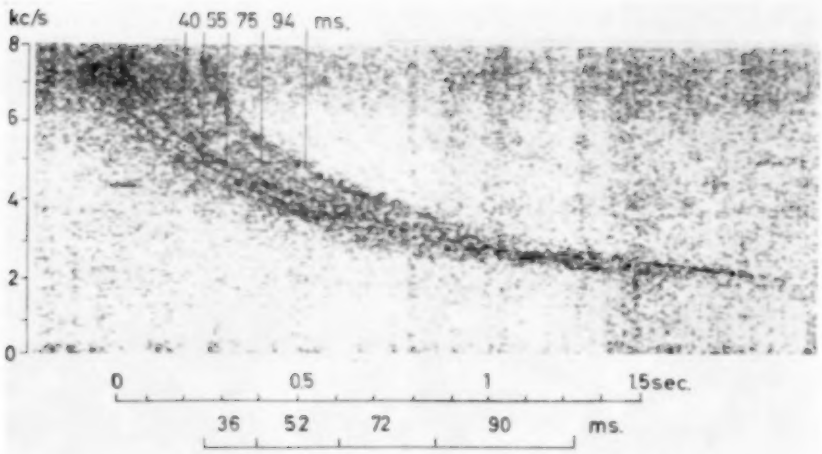


Fig. 8. Sonagram of five multiple whistlers with comparatively short time intervals.

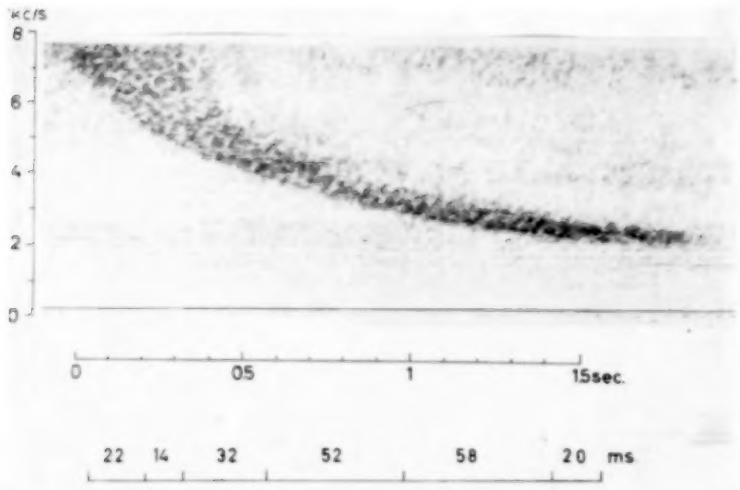


Fig. 9. Sonagram of a swish with six time intervals.

oscillographs was adapted to record time sequences and corresponding amplitude variations of multiple discharges in a lightning path whether followed by whistlers or not. In this way the necessary time differences were obtained on the "lightning side" of the problem.

On the "whistler side" a determination of the time intervals of multiple whistlers was obtained by a sound-spectrograph, the Sonagraph, manufactured by the Kay Electric Co. The sound-spectrograph produces records with very good details; one inconvenience is its slowness in recording. When time intervals of sonagrams are compared with corresponding ones from an oscillograph, the disagreement in time must be wholly ascribed to time uncertainties in the sonagram recordings. In the multiple whistlers illustrated the time intervals are calculated from the sonagrams and marked above the curves. Below the time-axes of the sonagrams are given the time intervals of multiple lightning discharges obtained by oscillograph.

In our earlier investigations it was found that a single discharge in a lightning channel occupies only 10–15 per cent of the total amount, the others being multiple strokes. A single whistler is a result of a single discharge in the lightning channel and must therefore be considered as a comparatively rare phenomenon. Such a single whistler is reproduced in Fig. 5, where the oscillographic analysis has recorded only one discharge within the total time period.

A sonagram which shows three multiple whistlers is reproduced in Fig. 6.

In Fig. 7 multiple whistlers are reproduced showing a first time interval of 366 m sec, which is unusually long compared with the following ones in the sequence.

A group of five multiple whistlers with comparatively short time intervals in the sequence is reproduced in Fig. 8. This accumulation of multiple whistlers mainly characterized by short intervals must be considered to be a " swish." A sonagram of such a type is reproduced in Fig. 9, where six intervals between lightning discharges in the same path follow each other. The intervals are short, which partially explains the peculiar variation form of a swish. On the other hand, it must be remarked that the intervals are not shorter than for other exemplified multiple discharges followed by distinct whistlers. A contributing effect for development of the specific variation form which a swish represents might be that the individual whistlers developed in this variation type are in some way dispersed and overlap each other in time.

From these comparisons of the time intervals of multiple whistlers obtained by oscillograms and by time–frequency sonagrams it is quite evident that multiple whistlers are caused by multiple discharges in the same lightning channel. The complete time agreement between the two phenomena has been demonstrated. The propagation of the multiple whistlers in the same lightning channels has also been confirmed by the agreement in the respective dispersion values of successive whistlers.

Acknowledgment—The geophysical research reported in this article has been sponsored and supported by Contract AF 61(052)-07, Electronics Research Directorate, Cambridge Research Center of the Air Research and Development Command, United States Air Force, through its European Office, Brussels.

REFERENCES

1. H. Norinder and E. Knudsen, *Planet. Space Sci.* **1**, 173 (1959).
2. H. Norinder, E. Knudsen and B. Vollmer, *Recent Advances in Atmospheric Electricity*, p. 525. Pergamon Press, London (1959).

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 50-58. Printed in Great Britain.

RADAR ECHOES OBTAINED FROM EARTH SATELLITES 1957 ALPHA AND 1957 BETA

W. E. JAYE, R. B. DYCE and R. L. LEADABRAND

Stanford Research Institute, Menlo Park, California

(*Received* 3 *April* 1960)

Abstract—On 16 occasions, radar echoes were obtained from the first Russian earth satellite, 1957 Alpha, by allowing the satellite to pass through the antenna beam, which was aimed at the point of closest approach.

Forty-eight echoes were obtained from 1957 Beta, the largest apparent cross-section being 437 m². A median cross-section of 1·8 m² was obtained. The echo exhibited a cyclic amplitude fluctuation which was observed to be as low as 0·1 c/s and as fast as 0·9 c/s.

Although calculations indicate that ionospheric Faraday rotation was present in many cases, the amplitude fluctuation of the back-scattered echo generally prevented its identification.

One occasion is reported of a rapid, irregular amplitude fluctuation imposed on the echo by aurora.

1. INTRODUCTION

As soon as the first satellite was launched, the interest in observing the effect of the ionosphere on coherent radio transmission was accompanied by attempts to obtain an echo from an object encircling the earth at high speed at altitudes in or above the *F*-region of the ionosphere. Success in obtaining such a radio echo was first reported by Stanford Research Institute (SRI)[1] and by Jodrell Bank Experimental Station.[2] As the present study was in progress, other successes were reported.[3, 4] The accomplishment of obtaining such echoes opened the possibility of exploring the transmission properties of the ionosphere using an object of known shape as a passive reflector. The echo strength and amplitude fluctuation characteristics were also of interest to determine whether or not ionization accompanied the satellite in its flight.

As part of a research program concerning aurora, meteor-trail and moon reflections, two experimental high-sensitivity radars were constructed and operated by SRI. The one finding greatest application for satellite observation operates at 106·1 Mc/s and is located at Stanford, California. The other radar, located at College, Alaska, operates at 398 Mc/s but has the same antenna dimensions and essentially the same parameters. Five observations were made on 398 Mc/s. It is sufficient to note that the results were in good agreement with the 106·1 Mc/s results discussed in this paper.

2. EQUIPMENT AND PROCEDURE

The data presented in this paper were obtained with a radar with the characteristics shown in Table 1.

Table 1. 106·1 Mc/s Radar Characteristics

Peak power	60 kW
Average power	4·5, 9 or 18 kW (depending on PRF selected)
Antenna gain	24 db (61-foot diameter steerable parabolic reflector, see Fig. 1)
Antenna beamwidth	12° between half-power points (one way)
Receiver noise figure	2·5 db
Receiver bandwidth	6 kc/s
Minimum detectable signal	approximately 10^{-16}
Minium detectable radar cross section at 500-km range	0·32 m²
Pulse width	1 m-sec
PRF	75, 150 or 300 c/s

The transmitted polarization is linearly vertical. Both vertical and horizontal linear polarizations are received.

The vertically polarized dipoles shown in Fig. 1 are fed in phase for symmetrical illumination of the dish. After this picture was taken, a similar pair of horizontal dipoles were added between the end of the vertical dipoles for cross-polarized reception. The signal is fed to two independent preamplifier–converters and from there into two Collins 51J4 receivers having a bandwidth of 6 kc/s. The noise figure of the receiving system, including feed and T–R switches, is less than 5 db.

An *A*-scan oscilloscope presentation is available for on the spot observations and single-shot photography. Most of the analysis was made with the convenience of magnetic tape recording, one channel for each of the two received polarizations.

At the time of these measurements, accurate predictions of the satellite motion were lacking. Early orbit prediction was done by measuring the maximum rate of change of the Doppler frequency shift of the satellite-emitted signal.[5] This was sufficient to provide day to day radar-pointing data. This method was abandoned when Project Space Track of the Geophysics Research Directorate at Air Force Cambridge Research Center began to issue bulletins containing acquisition data.

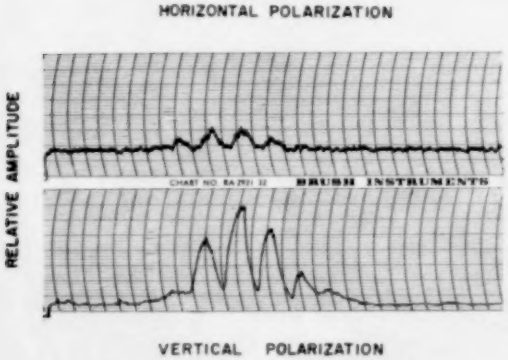


Fig. 2. A typical amplitude-vs.-time pattern of an echo observed from satellite 1957 Beta. 0605: 30.9 March 1958 U.T. Slant range = 340 km. azimuth = 120°, elevation = 52·5°, 1 scale division = 1sec.

Owing to the lack of knowledge of the real time position and motions of the satellite, and to minimize the range and requirement for Doppler shift correction, the antenna was generally fixed in the direction of closest approach (at right angles to the satellite's trajectory). If the elevation angle setting was correct, the satellite echo was observed to gradually rise out of the background noise, fluctuate several times (due to tumbling) and gradually sink back into the noise again, the entire process taking on the order of 10 sec (see Figs. 2 and 3). The recorded signal strengths in the two polarization channels were mathematically combined to yield the total intensity and the apparent received polarization (which unfortunately has an unresolvable 90° ambiguity). The example shown in Fig. 2 has been treated in this way to yield the curves of Fig. 3.

The time of occurrence of maximum echo envelope was taken to be the time that the satellite crossed the azimuth in which the antenna was pointed (generally 60°, 120°, 240° or 300° true). Time of passage through the central azimuth beam can be determined to an accuracy of a few seconds of time, largely limited by the broadness of the antenna beam. The radar range accuracy is determined primarily by the receiver rise time and is about ±10 km when a sufficiently strong echo is present.

The time and range of these events, the observed echo equivalent cross-section and the nature of the echo fluctuations while within the beam are discussed in the following sections.

3. OBSERVATIONS

One of the most important problems was the discrimination between echoes from meteor trail ionization and the echo from the satellite. At 106·1 Mc/s, meteor-trail echoes usually have a very fast rise-time, on the order of a few milliseconds. Duration of this echo depends on the size and speed of the meteor, and is usually less than 1 sec. (A rapidly tumbling satellite, such as 1958 Delta-1, may also have an echo rise-time almost as fast.) Occasionally, a very large meteor with an over-dense trail will give an echo resembling a satellite. These echoes may last a long time (more than 5 sec), have a slower rise time than ordinary meteor trail echoes and

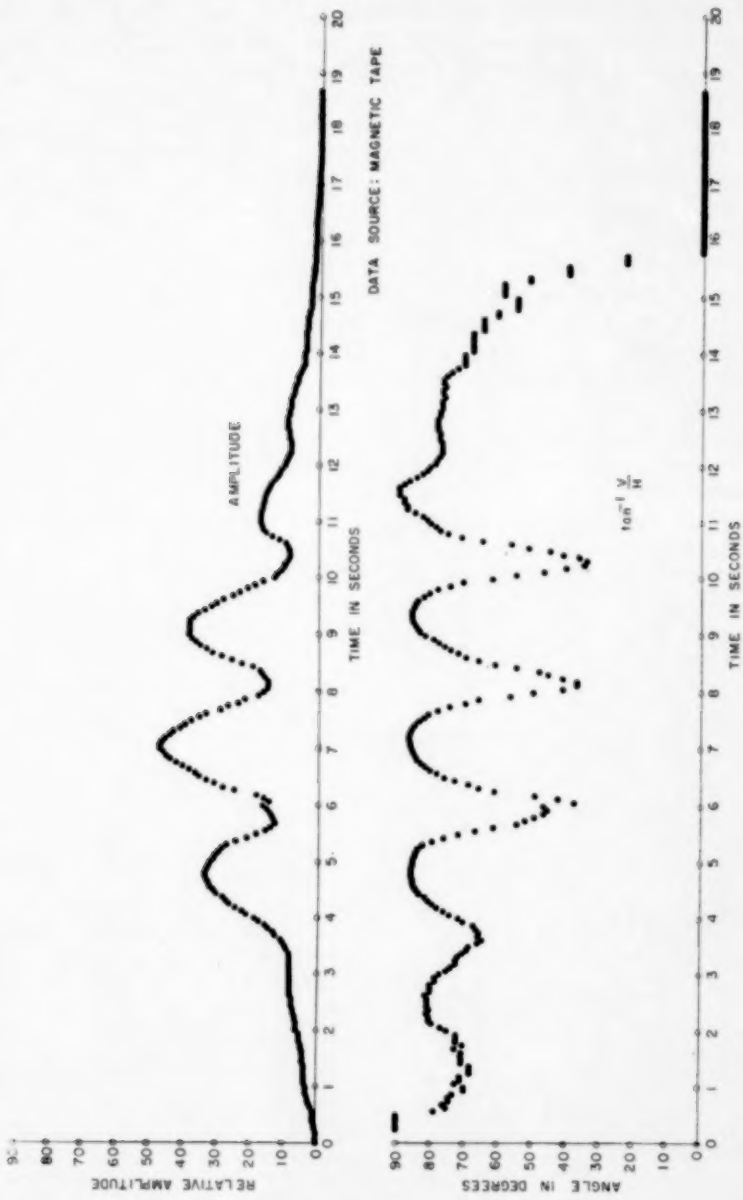


Fig. 3. Total echo amplitude and polarization angle vs. time for echo shown in Fig. 2.

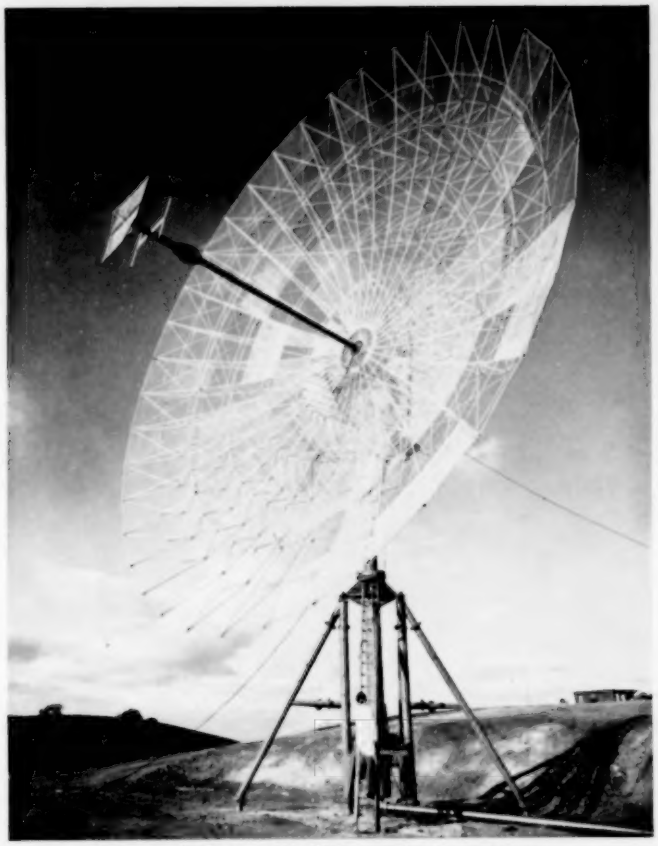


Fig. 1. 61-foot steerable parabolic reflector.

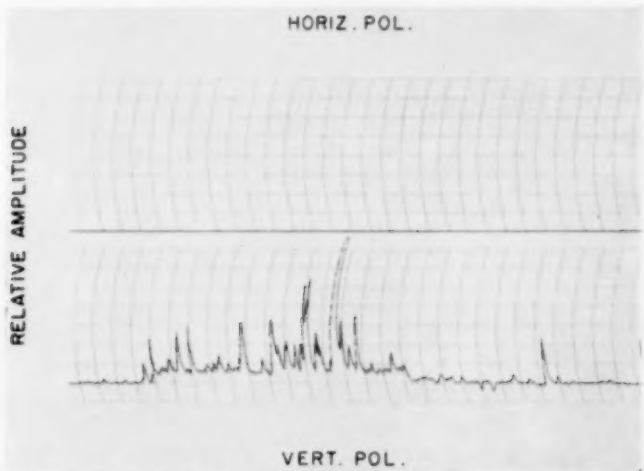


Fig. 8. Echo from 1957 Beta under auroral conditions. 0454.16 March 1958 U.T., slant range = 700 km, azimuth = 300°, elevation = 24°. (height = 282 km). 1 scale division = 1 sec.

show cyclic amplitude fluctuations. Thus, their echo characteristics may occasionally resemble satellite echoes.

Because meteor echoes generally occur between 80 and 120 km above the surface of the earth and because satellite echoes are normally obtained at ranges exceeding 200 km, discrimination would seem possible, based on the beam elevation angle and range. Unfortunately, a strong meteor echo can be observed in the sidelobes of the antenna. These echoes occur at ranges similar to the range of the expected satellite echoes, making discrimination solely on the basis of range impossible.

At long ranges, these satellites could stay in the 12° beam for 20 to 25 sec. It is highly unlikely that a meteor-trail echo will occur at the calculated satellite range at the time that a satellite is expected, and having a duration as long as the expected satellite echo.

Identification of the echo from meteor-trail echoes and from background noise is aided greatly by the fact that a cyclic amplitude pattern is generally present. Differentiation of the echo from the background noise is roughly proportional to the number of echo oscillations present.

Sixteen echoes were obtained from 1957 Alpha. Forty-eight echoes were obtained from 1957 Beta (average duration of 13·9 sec).

4. ANALYSIS OF DATA

Satellite 1957 Beta was tumbling about an axis almost perpendicular to the spin axis[3] with a period of about 200 sec. Spin about the axis of symmetry of the satellite was not known. Gradual motion of the total angular momentum axis was also present. The tumbling causes the aspect of the satellite to change rapidly with respect to the radio ray during the transit through the radio beam. A large number of different cross-sections and amplitude fluctuation patterns are therefore expected which may vary considerably from day to day.

A. Echo cross-section

The apparent radar cross-section can be calculated from the following equation:

$$\sigma = \sigma_{\min} (S/N)^2 (R/500)^4 (1/g_p)^4 \qquad (1)$$

where

σ = apparent radar cross-section (m^2)

$\sigma_{\min}$ = 0·32 m^2 at 500 km

S = signal amplitude

N = noise amplitude

R = slant range (km)

g_p = normalized field strength in a given direction ($g_p = 1$, for center of beam)

The minimum detectable signal at 500 km was calculated assuming a transmitter power of 60 kW, a bandwidth of 6 kc/s, an ambient temperature of 20°C, an antenna gain of 24 db, and a system loss of 5 db.

The histogram of Fig. 4 illustrates the statistics of the observed apparent cross-section. Fifty per cent of the echoes obtained were smaller than 1·8 m^2.

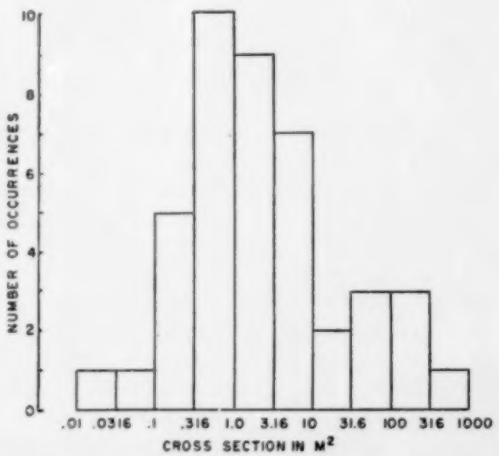


Fig. 4. 1957 Beta echo occurrences vs. apparent cross section.

The largest cross-section observed was 437 m^2. Let us assume that this value represents the maximum possible cross-section obtainable from a satellite of cylindrical shape. The cross-section of a cylinder at normal incidence (maximum cross-section) is ideally given by:

$$\sigma = \frac{2\pi L^2 a}{\lambda} \qquad (2)$$

where

L = length of the cylinder

a = radius of the cylinder

λ = wavelength

For a length to diameter ratio of 8, for example, this is consistent with a cylinder about 15 m long and 2 m in diameter.

B. Amplitude fluctuation

As the satellite crosses the beam, a periodic variation of signal amplitude is usually apparent because of the variation of cross-section with aspect angle. The period of these fades was observed to be as low as 0·1 or as high as 0·9 c/s. A typical example of lobe structure as the satellite traverses the beam was shown in Fig. 2.

C. Polarization

Usually the periodic lobes of the echo observed in one polarization and the lobes of the echo observed in the orthogonal polarization are in agreement as though the total power being reflected was varying in periodic manner, unaccompanied by large changes of polarization. In some instances, however, the phase relationship of the oscillations in the two polarization channels slowly changed (see Figs. 5 and 6). Initially some of this change was suspected to be caused by Faraday rotation in the ionosphere.

The plane of polarization is known to be rotated during the double passage of the 106·1 Mc/s wave through the entire ionosphere.[6] Ionosonde data taken at Stanford University was used for calculating an electron density profile in the manner of Budden.[7] Typical results are given in Fig. 7. In some cases, the calculated polarization twist due to the ionosphere was small or zero (when the satellite was below the densest part of the ionosphere). On other occasions, when the electron content between the bottom of the ionospheric layer and the satellite was large, the two-way polarization rotation could have been as great as three rotations. The example shown in Fig. 2 was made under conditions of no significant polarization rotation. (Ionospheric conditions shown in Fig. 7 were present.) However, large polarization twist apparently has no profound effect on the echo amplitude behaviour. Unfortunately, dual-polarization data are not available for passes during which a high polarization twist was calculated.

The amount of Faraday rotation is largely a function of the path length through the ionosphere (dependent upon the angle of incidence) and the angle between the radio ray and the magnetic intensity vector. Both of these angles change as the satellite traverses the radar beam.

The combination of these two effects may cause the two-way Faraday rotation to change by as much as 25 per cent at low elevation angles. Unfortunately, the effect is not clearly demonstrated due to the lack of dual-polarization data on those passes having a significant change of Faraday rotation during a pass. Several passes were available (both polarization channels functioning) when the calculated Faraday rotation exceeded 21 degrees. As anticipated, no observable effects on the echo were apparent. The apparent changes in polarization demonstrated by Fig. 6 are not explainable by Faraday rotation, since no significant Faraday rotation or change of rotation was present during this occasion. One is left with the conclusion that such cross-section and polarization changes are a consequence of the satellite shape and tumbling motion alone.

D. Effect of aurora

On one occasion, satellite 1957 Beta was detected in the vicinity of an auroral disturbance. Prior to the passage of the satellite, auroral echoes were located approximately 10° farther north in azimuth, and 200 km farther in slant range. (No auroral echo was present when the satellite was detected because the condition of orthogonality required for auroral reflection was not satisfied at the azimuth of satellite passage.) On this pass, the echo was subject to unusually rapid and irregular fluctuations (Fig. 8) apparently due to the aurora. The signal endured longer than a typical echo but was not significantly stronger or weaker.

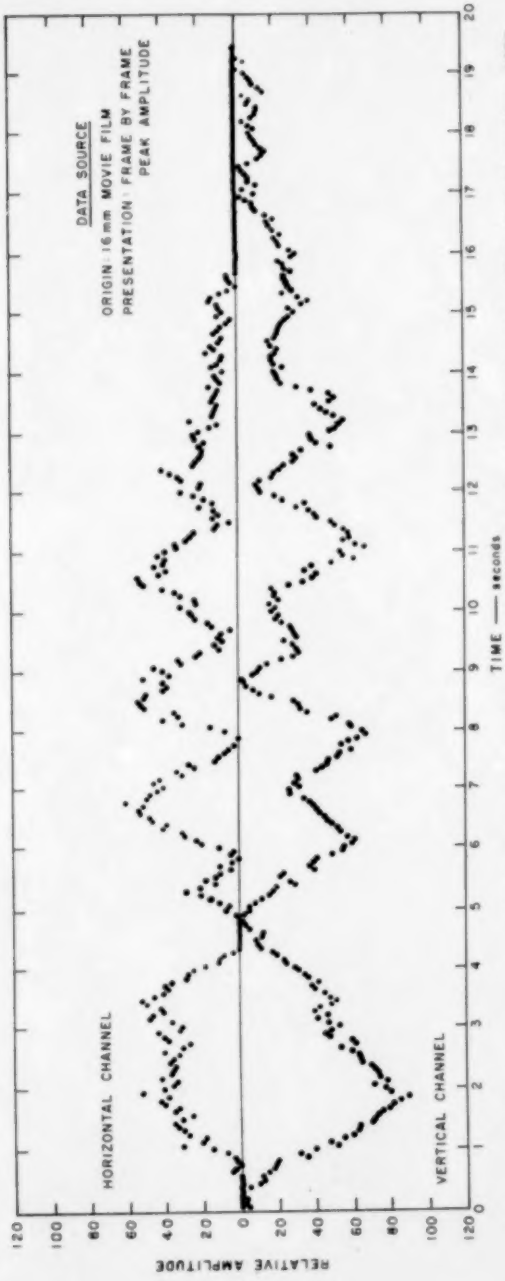


Fig. 5. Amplitude vs. time of 1957 Beta echo, showing uncorrelated fluctuation in each channel. 1338: 05.7 November 1957 U.T. Slant range = 420 km, azimuth = 300°, elevation = 32°.

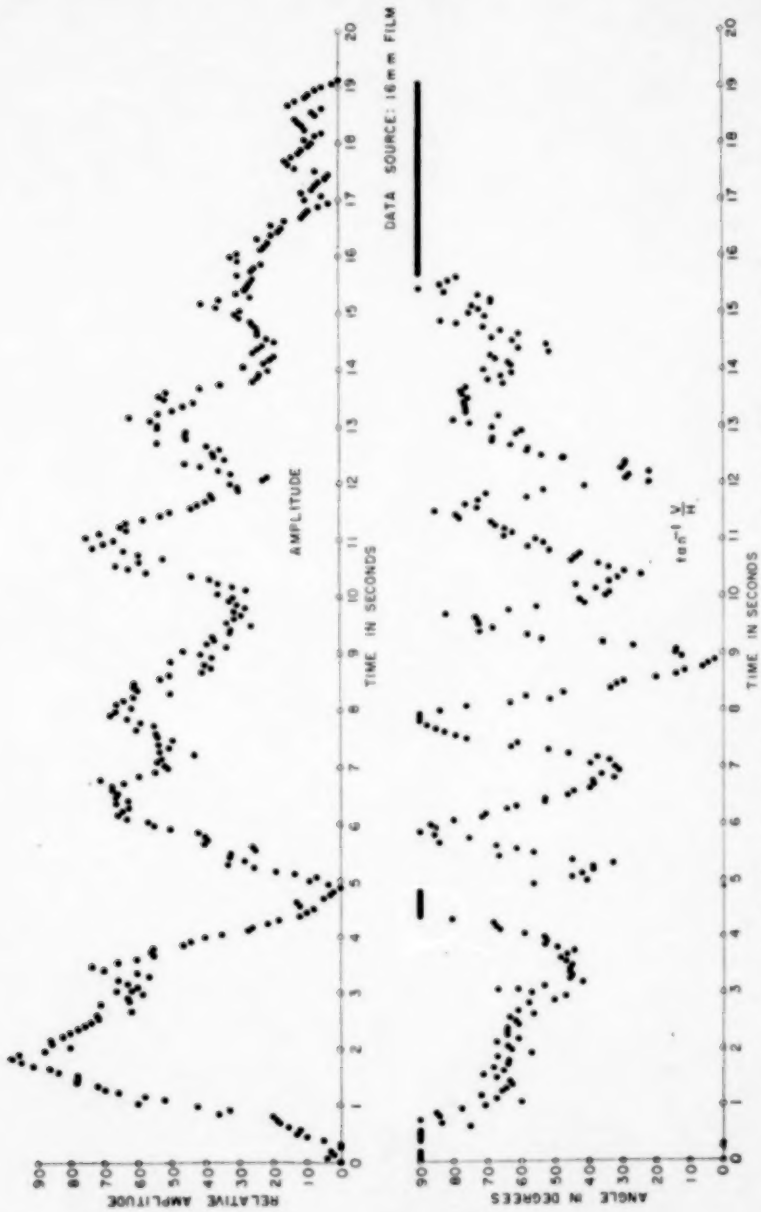


Fig. 6. Total amplitude and polarization angle vs. time for echo shown in Fig. 5.

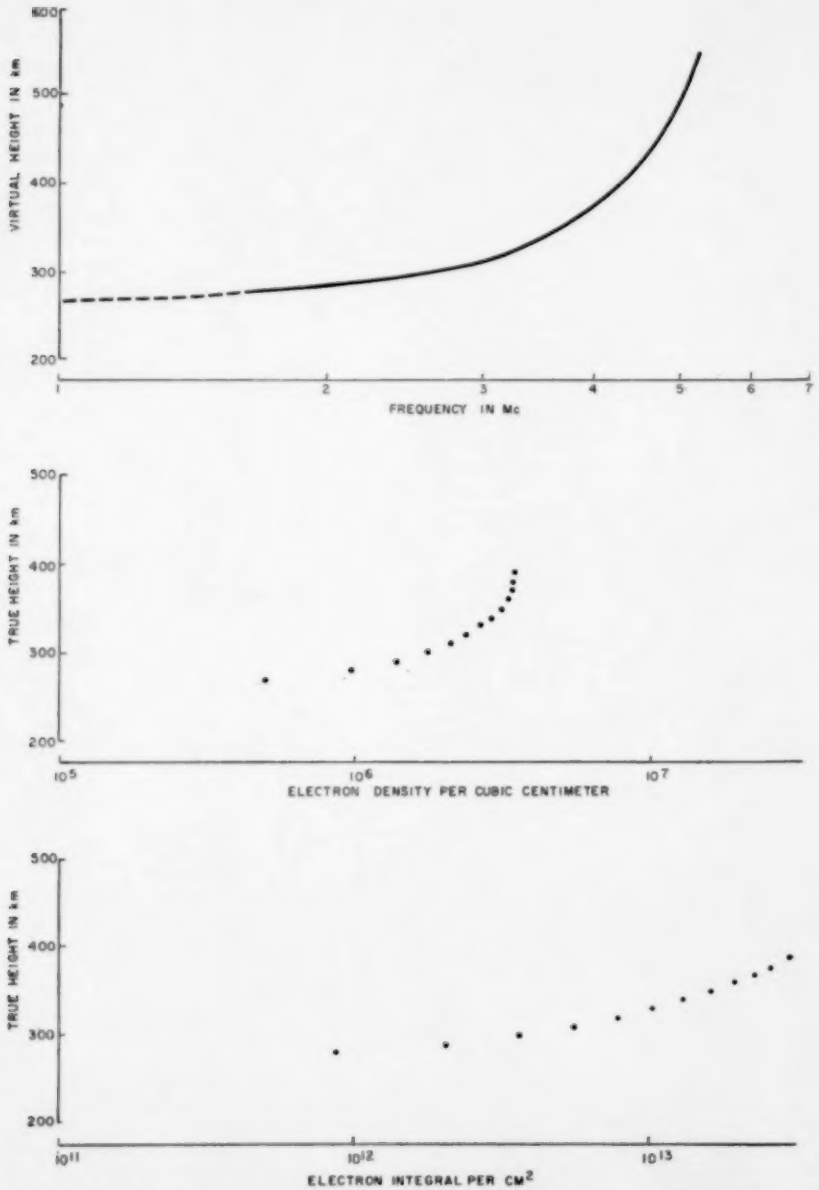


Fig. 7. (a) Ionogram, (b) electron density profile content and (c) integrated electron content for Stanford, California, 0615.9 March 1958 U.T.

5. CONCLUSIONS AND RECOMMENDATIONS

Radar echoes at 106·1 Mc/s from satellite 1957 Beta reveal a median cross-section of 1·8 m^2, based on forty-two observations. Apparently there is no enhancement of the echo due to a large trail of orbiting ionization. The most striking characteristic of the echo is the cyclic amplitude variation caused by satellite tumble. In principle, knowing the rate of rotation of the satellite, the geometry of the satellite might be deduced.

Some curious polarization behaviour is observed. Faraday polarization rotation in the ionosphere is not adequate to explain this phenomenon. This polarization change is attributed to the reflection behavior of the satellite itself.

One occasion is reported of the possible effect of aurora on the transmission of radio waves to and from the satellite.

The technique of using the earth satellite as a passive radar reflector for the purpose of ionospheric studies could well be augmented in at least two ways:

(a) The use of a spherical satellite without protruding antennas,

(b) The use of high-inclination orbits so that propagation through the auroral ionization might be examined.

Acknowledgments—The authors express thanks to many persons within the Laboratory for their faithful assistance in operation of the radar. Also, the efforts of L. H. Rorden were vital to the early orbit determinations. This work was made possible by the support of Rome Air Development Center of the Air Research and Development Command of the U.S. Air Force under Contract AF 30(602)-1762.

REFERENCES

1. Special Techniques Group, SRI, *Proc. Inst. Radio. Engrs.* **45**, 11 (1957).
2. Jodrell Bank Experimental Station, *Nature, Lond.* **180**, 4593 (1957).
3. D. K. Barton, *Sputnik II as Observed by C-Band Radar* RCA, Part V, p. 67, National Convention Record, IRE (1959).
4. G. H. Pettengill and L. G. Kraft, Jr., *Earth Satellite Observations Made with Millstone Hill Radar* Lincoln Laboritories, MIT, Paper given at AGARD Avionics Panel Meeting, Copenhagen, Denmark (1958).
5. W. E. Jaye and L. H. Rorden, *Final Rep.* Part III, SRI, Contr. AF 30(602)-1762 (1959).
6. C. G. Little, W. M. Rayton and R. B. Roof, *Proc. Inst. Radio Engrs.* **44**, 8 (1956).
7. K. G. Budden, *Physics of the Ionosphere*, pp.276-87. Physical Society, London (1955).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 59-69. Printed in Great Britain.

POLAR IONOSPHERIC DISTURBANCES AND SOLAR CORPUSCULAR EMISSIONS

T. OBAYASHI and Y. HAKURA

Radio Research Laboratories, Kokubunji, Tokyo

(*Received* 3 *November* 1959)

Abstract—It has been found that the study of polar radio blackouts due to abnormal ionization in the lower ionosphere yields considerable evidence indicating the existence of energetic solar particles associated with solar flares. Polar radio blackouts are classified into two characteristic types, one is the polar cap blackout and the other is the auroral zone blackout. It is shown that the polar cap blackout appears with some hours delay after a major solar radio outburst of type IV, and the blackout is confined within the geomagnetic latitude of 60°–65°. The estimated energies of particles causing this are of about 10–100 MeV. The auroral zone blackout then follows, being accompanied with geomagnetic storms and aurorae, and it may be caused by the so-called auroral particles of 1 MeV or less.

The energy spectrum of solar particles associated with solar flares is revealed from the present result together with all information from various observations related to solar and terrestrial disturbances. It is concluded that solar particles have a conspicuous suprathermal non-Maxwellian tail extending from a few keV up to relativistic energy range, though the bulk of corpuscular clouds consists of rather low energy particles and hence likely to be in the Maxwellian distribution. Some discussions on the nature of solar corpuscular clouds and their effect upon the terrestrial ionosphere are also given.

1. INTRODUCTION

One of the important projects aimed at in the Third International Geophysical Year 1957–1958 was to establish a better understanding of the physical aspects of disturbances in the earth's upper atmosphere. Since the greatest solar activity was attained during this period, considerable interest has been stimulated on the research concerning solar disturbances and their effects upon the terrestrial ionosphere. Some new results are forthcoming, and comprehensive studies are now being concentrated to solve this attractive problem.

For the study of "earth storms", the transient radiative and corpuscular emissions associated with solar flares are the most important and prominent subjects. Solar flares produce the excess of ultraviolet radiation which extends into the X-ray range, causing the effect so-called sudden ionospheric disturbances. It is generally believed that the corpuscular emissions are followed by the ionized jet plasma travelling with velocities of the order of 1000 km/sec, which would produce geomagnetic and ionospheric storms. At the time of a great solar flare, the sun also ejects corpuscles of relativistic energies, and occasionally they can be observed as an abrupt increase of the cosmic ray intensity, though only a few cases are known until at present.

Recently, Bailey[1] suggested the existence of new solar particles in the sub-cosmic ray energy range from his study of the polar ionospheric absorptions during the great solar event of 23 February 1956. From their analysis of the IGY data, the present authors[2, 3] have also found considerable evidence indicating that energetic charged particles of 10–100 MeV reach the earth with some delay from the solar flare, yet well before the onset of a geomagnetic storm. They impinge on the lower ionosphere at high latitudes and cause the polar cap blackouts. During geomagnetic storms, however, the particles precipitate along the auroral zone producing aurorae and also the auroral zone blackouts. Energy of these so-called auroral

particles is presumably lower, being of about 1 MeV or less. Independent confirmations of the existence of these new particles have been given by the discovery of abnormal absorptions of cosmic radio noise observed in the polar cap regions,[4, 5] also by floating balloon observations of cosmic rays at high latitudes[6] and by satellite, Pioneer IV.[7]

Since these evidences may provide a new way to investigate the nature of solar corpuscular emissions and the mechanism of geomagnetic storms, further studies are very promising. In the present paper, the characteristics of enhanced ionizations in the polar ionosphere associated with solar flares are briefly reviewed. The statistical relationship between solar and terrestrial disturbances is re-examined, and it is confirmed that major radio outbursts of type IV are closely related to the polar cap blackouts preceding geomagnetic storms. From these results together with all other information from various observations related to solar and terrestrial disturbances, a tentative model of the energy spectrum of solar particles is proposed. It is shown that the ejected particles associated with solar flares have a considerably wide energy spectrum extending up to relativistic range, and that these particles may mostly be generated in the solar corona during flares by the Fermi acceleration mechanism. The effect of invading solar particles upon the upper atmosphere and also the mechanism of earth storms are discussed briefly.

2. TERRESTRIAL DISTURBANCES ASSOCIATED WITH SOLAR FLARES

It is well known that, at the solar flare, an outstanding radio emission from the solar corona, which is called the solar radio outburst, is observed. An investigation of dynamic spectra of radio outbursts reveals some important nature of the disturbance in the solar atmosphere. As has been pointed out by one of the authors[8], intense solar radio outbursts, whose maximum power spectrum lies in the SHF band, cause sudden ionospheric disturbances. They are explained as the effect of the enhanced ionization in the ionosphere due to excess of the soft X-ray radiation. On the other hand, major radio outbursts predominant in UHF and VHF bands, especially those of continuum radiation named type IV,[9] are closely correlated with severe geomagnetic storms.

It has been revealed, recently, that an enhanced ionization in the ionosphere, which is indicated by the increased f_{min} or the radio blackout of echoes in the ionograms, occurs in

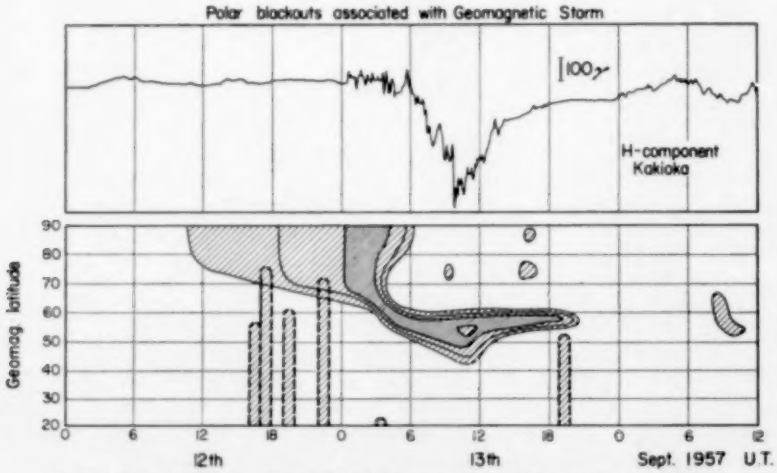


Fig. 1. Latitude effect of polar blackouts associated with a geomagnetic storm of 13, Sept. 1957 (hatched region shows abnormal increase of f_{min}).

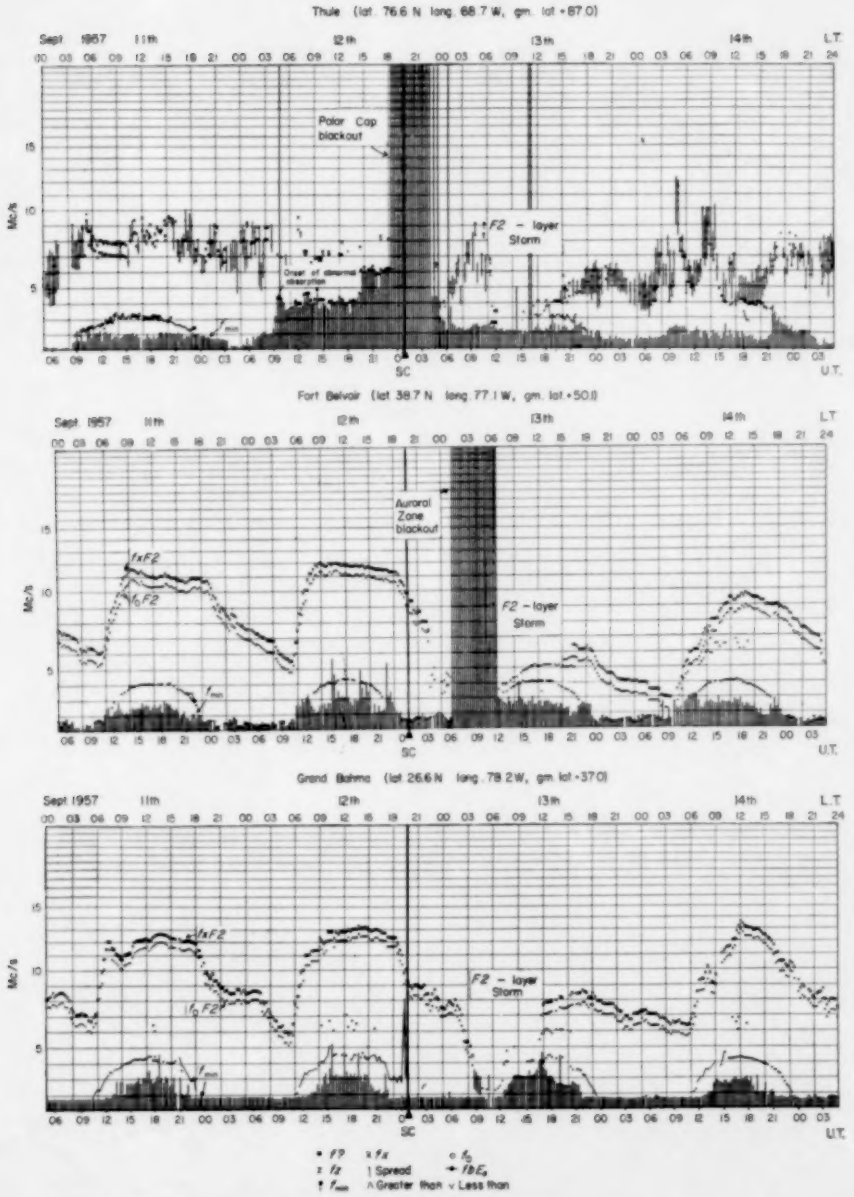


Fig. 2. *f*-plots of ionospheric vertical soundings on Sept. 11–14, 1957, at Thule (polar cap), Fort Belvoir (sub-auroral zone) and Grand Bahma (middle latitude).

the polar region several hours after a major radio outburst of type IV and, yet, well before the onset of a geomagnetic storm.[2] Since it is thought that polar radio blackouts are caused by the abnormal ionizations due to bombardment of energetic particles, this fact would suggest the invasion of solar particles presumably of higher energies than those producing geomagnetic storms. In order to show the general aspect of this phenomenon, a typical one observed on 12–13 September 1957 is described. A major type-IV solar radio outburst was observed at 2 hr 45 min U.T. on 11 September, associated with an intense solar flare. The variation of f_{min} vs. geomagnetic latitude following this solar event is illustrated in Fig. 1, using data of about twenty ionospheric stations in the North America Zone and with the geomagnetic record at Kakioka, Japan. The hatched regions indicate increased f_{min} or complete blackouts (dense hatching). The regions enclosed by dotted lines are those of increased f_{min} caused by SID's and they are confined in the middle and low latitudes of the sunlit hemisphere.

An outstanding increase of f_{min} appeared at high latitudes more than 10 hr before the onset of geomagnetic storm. Intense polar blackouts were then observed along the auroral zones in the earlier stages of the storm, while in the last phase any systematic wide-spread blackout was not observed. The original f-plots obtained at Thule (geomagnetic latitude 87·0°), Fort Belvoir (50·1°) and Grand Bahma (37·0°) are also reproduced in Fig. 2. It is noted that an abrupt increase of f_{min} was noticed at 9 hr U.T. on 12 September at Thule (polar cap station), while no appreciable change appeared at Fort Belvoir (sub-auroral zone station) until at the main phase of the geomagnetic storm.

The world-wide patterns of the abnormal ionization for respective stages of geomagnetic storm are characteristic. As shown in Fig. 3, the one which appears before the onset of geomagnetic storms is that the enhanced ionization (radio blackouts) is confined within the polar cap, while during geomagnetic storms it spreads along the auroral zone forming a spiral shaped ionization pattern. The former is termed the *polar cap blackout* and apparently no marked magnetic activity is associated, while the latter is the *auroral zone blackout* which is for the most part a night-time effect and its occurrence is correlated with that of visible aurorae and local geomagnetic disturbances. It has been shown elsewhere[3] that the polar cap blackout

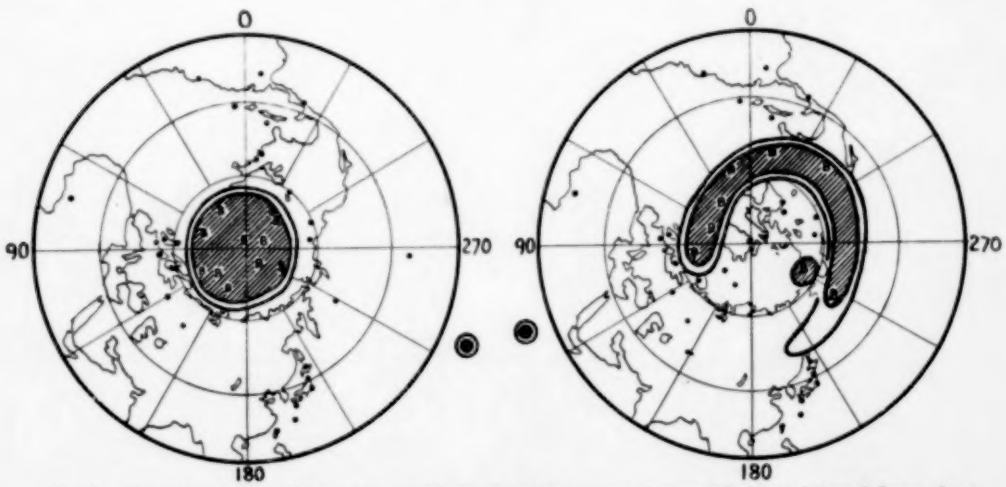


Fig. 3. Typical patterns of the polar cap blackout and the auroral zone blackout (viewed from above the geomagnetic north pole) at the stage before the onset of the geogmagnetic storm (pre-s.c.), 00 hr U.T., and the main phase, 10 hr U.T., on 13 Sept., 1957.

is caused by the invasion of charged particles of about 10–100 MeV and those particles would impinge on the polar ionosphere confined within the geomagnetic latitude of 60°–65°, providing that they follow the Störmerian orbits. The auroral zone blackout is related to auroral displays and it is presumably due to auroral particles of 1 MeV or less. It is noted that the precipitation of these auroral particles on the earth is strongly affected by the effect of distortion of the outer geomagnetic field during geomagnetic storms.

As has already been emphasized, it is very important that major solar radio outbursts are closely related to those disturbances which indicate the invasion of energetic particles into the earth. In order to confirm this, a further statistical analysis is made using the IGY data of solar radiometry, polar blackouts and geomagnetism. During the period July 1957 to December 1958, there were fifty-nine s.s.c. magnetic storms, among them twenty were followed by major radio outbursts with type IV. It has been found that prolonged severe *polar cap blackouts* were always associated with those of the type-IV outbursts except three rather doubtful cases. The occurrence probability of the polar cap blackouts following type IV outbursts is about 75 per cent, while without type IV it is only 8 per cent. The result is summarized in Table 1, and more details of this analysis is given elsewhere.[9,10]

Table 1. Occurrence of Magnetic Storms and Polar Cap Blackouts

Sun	Magnetic storms	Polar cap blackouts
Major outburst with type IV	20	15
Major outburst without type IV	39	3

Since a type-IV outburst is caused by synchrotron radiation due to relativistic electrons spiralling in magnetic fields, the generation of high-energy particles in the solar corona would likely be expected. A dynamical process creating "suprathermal particles" associated with solar flares has been suggested by Hayakawa and Kitao[11] and by Parker.[12] They showed that an agitated solar plasma bearing magnetic fields may be capable of accelerating protons from thermal to relativistic energies by Fermi's mechanism. Therefore corpuscular clouds ejected from the sun may contain a considerable amount of energetic protons, and they will arrive at the earth with some delay after the type-IV outburst.

It is likely that ejected clouds would trap the magnetic fields of sunspots or the corona and travel outward through the solar atmosphere into interplanetary space. High-energy particles may be partly trapped in such magnetic clouds during the earlier stage of their passage through interplanetary space. However, as the cloud undergoes an expansion and accordingly magnetic fields of the cloud dilute, $\overline{H}^2 \sim n^{4/3}$, the high-energy particles may gradually escape from the cloud and reach the earth causing *polar cap blackouts* before the time of arrival of the cloud, which is indicated by the onset of a magnetic storm. This implies that high-energy particles may not come directly straight from the sun to the earth, and the observed delay of the polar cap blackouts after solar flares at least of the order of a few hours is explainable on this basis. Although it has been a puzzling problem that the incoming speed of auroral particles estimated from the observations is several times the speed of corpuscular clouds inferred from solar terrestrial relationships, it can now be understood that a considerable amount of high-energy particles exists in the corpuscular cloud, which could provide the energy of auroral displays or polar blackouts, even though the speed of the cloud itself is only of the order of 1000 km/sec.

Summarizing the present results together with other knowledge presently at hand, a general scheme of solar and terrestrial disturbances is depicted in Table 2.

High-energy protons of the order of 10–100 MeV reach the earth several hours after a solar flare accompanied with major radio out-

Table 2. Solar Flares and Disturbances in the Earth's Upper Atmosphere

Flare	*e.m. Radiations*	Soft X-rays H_α Enhancement and Major Radio Outbursts	SID and Crochet
Corpuscular cloud	*Relativistic electrons* ...	Synchrotron radiation	Type IV outbursts
	Relativistic protons ...	Solar cosmic ray	Unusual increase of cosmic ray
	High energy particles ...	Solar cosmic ray	Polar cap blackouts
		Auroral particles	Auroral zone blackouts
		(trapped in the cloud)	Aurorae, polar magnetic storms
	Low energy particles ...	Magnetic cloud	
		Corpuscular cavity formation ...	Dst of magnetic storms
		Magnetic barrier for cosmic rays	Forbush-type decrease

bursts of type IV, while occasionally solar cosmic ray particles well above relativistic energies are observed associated with particular great solar flares.[12] Those high-energy particles invade into the polar regions following nearly Störmer orbits and cause prolonged *polar cap blackouts* or abnormal VHF absorption at high latitudes, so-called type III absorption.[4, 5] On the other hand, the bulk of a corpuscular cloud consisting mainly of low-energy particles arrives after a day or so, corresponding to a speed of the order of 1000 km/sec. The particle density of the cloud is considerably high and hence the cloud behaves like a conducting fluid. The interaction of such a conducting cloud with the geomagnetic field is exactly the same as reported by Chapman and Ferraro.[14] As has been pointed out by one of the authors,[15] the geomagnetic field is confined within a cavity surrounding the earth, and the outside of this is exposed to violently agitated solar plasma. The size of the cavity is variable according to the kinetic energy density of the incoming corpuscular cloud, and the storm-time variation Dst of geomagnetic storms may in some way be related to this cavity formation. It is noted that auroral particles, trapped by magnetic fields (order of 10^{-3}G) of the corpuscular cloud, reach the earth after the arrival of the cloud. They impinge on the polar ionosphere through the distorted geomagnetic field and cause *auroral zone blackouts*, aurorae and polar magnetic storms. An equatorward shift of the auroral zone during the main phase of magnetic storms is explained by the distortion of the geomagnetic field due to the change of the geomagnetic cavity.[3] The corpuscular clouds advance further into outer space until to a distance of a few astronomical units, where they form magnetic barriers for cosmic ray particles coming from the interstellar space.[13, 15]

3. THE ENERGY SPECTRUM OF SOLAR PARTICLES

Although it has been shown in the preceding section that the sun radiates an enormous amount of corpuscles—ranging from thermal to relativistic energies—in order to have further insight into the problem, it is of special importance to know the energy spectrum of particles ejected from the sun. Together with the present result, some other information is now available, and it is possible to delineate the energy spectrum of particles associated with solar flares at least consistent with some principal facts of terrestrial disturbances.

To unify the data of various observations into the same scale, every information is converted in terms of the integral energy spectrum, i.e. the particle flux $\mathbf{J}$ cm^{-2}sec^{-1} having energies in excess of the specified value E is defined as

$$\mathbf{J}(>E)=\int_{E}^{\infty}\mathbf{j}(E)\,dE \qquad (1)$$

where **j** is the differential energy spectrum. Results obtained by various investigators are summarized in Table 3, in which some data are given in terms of magnetic rigidity instead of energy. Magnetic rigidity **P** is given as

$$\mathbf{P} = \frac{p}{z} = \frac{1}{ze}\sqrt{E^2 + 2m_0c^2E} \qquad (2)$$

where ze is the charge (e.s.u.) and m_0c^2 is a rest energy.

particles. The abnormal ionization in the polar ionosphere began gradually and reached a maximum a few hours after the flare. Bailey[1] showed that the energetic protons of 10–1000 MeV are responsible for this. His estimation (2) is based mainly on the results of balloon observations by Van Allen and Winckler, which show the rigidity spectrum varying with $\mathbf{P}^{-5}$ at a range of 1000 MeV. The spectrum in the energy range between 1 and 100 MeV

Table 3. List of Solar Particles

	Energy range	Integral spectrum	Investigators	Year	Remarks
Solar cosmic ray	2–15 BV	$J_0(P/P_0)^{-7}$	Meyer, Parker, Simpson	1956	5 hr Feb. 23 1956
	0·5 BV~	$J_0(P/P_0)^{-5}$	Van Allen, Winckler	1957	21 hr Feb. 23 1956
		$J_0 = 1{\cdot}3/\text{cm}^2/\text{sec-sterad}$. $P_0 = 1{\cdot}2$ BV			(Balloon Obs.)
	0·5 BV	$J_0(P/P_0)^{-1}$	Bailey	1959	21 hr Feb. 23 1956
	100–400 MeV	$J_0 = (E/E_0)^{-4}$	Winckler *et al*	1959	11 hr Aug. 23 1958
		$J_0 = 1{\cdot}5/\text{cm}^2/\text{sec}$, $E_0 = 100$ MeV			(Balloon Obs.)
	30–100 MeV	$J_0 \sim 500/\text{cm}^2/\text{sec}$	Rothwell, McIlwain	1959	Satellites, Aug. 1958
	10 MeV	$J_0 \sim 10^2/\text{cm}^2/\text{sec}$	Hakura, Obayashi	1959	Polar cap blackouts
	10 MeV	$J_0 \simeq 0{\cdot}8 \times 10^2/\text{cm}^2/\text{sec}$	Reid, Collins	1959	Type III absorption
Auroral particles	10–100 KeV	$J_0 = 10^7$–$10^8\text{cm}^2/\text{sec}$	Chamberlain	1954	Auroral doppler shift
	30–500 KeV	$J_0 \exp(-E/E_0)$	Meredith *et al*	1958	Rocket obs.
		$J_0 = 1{\cdot}8 \times 10^5/\text{cm}^2/\text{sec-sterad}$. $E_0 = 72$ KeV			(Ion counter)
	10–100 KeV	$J_0 \sim 10^6$–$10^8/\text{cm}^2/\text{sec}$	Van Allen	1958	Ion electrons
	1 MeV	$J_0 \sim 10^5/\text{cm}^2/\text{sec}$	Hakura, Obayashi	1959	Auroral zone blackouts
Corpuscles	2000 km/sec	$J_0 = 2 \times 10^{10}/\text{cm}^2/\text{sec}$	Chapman	1950	Magnetic storms
	500–1500 km/sec	$J_0 \sim 10^{10}$–$10^{12}/\text{cm}^2/\text{sec}$	Biermann	1957	Comet tail
	10 eV–1 KeV	$J_0 \sim 10^9$–$10^{12}/\text{cm}^2/\text{sec}$	Obayashi	1958	Storm pulsations
Others	2–15 BV	$J_0(P/P_0)^{-1\cdot1}$	Van Allen, Singer	1950	Cosmic rays
		$J_0 = 0{\cdot}5/\text{cm}^2/\text{sec-sterad}$. $P_0 = 1$ BV			
	30 eV	$J_0 = \frac{1}{4}\,nv10^9/\text{cm}^2/\text{sec}$	Chapman	1957	Extension of corona

Combining all the results, the integral energy spectrum of solar particles is illustrated in Fig. 4. The information of solar particles of the highest energy is given by the unusual increase of cosmic ray intensity associated with the great solar flare on 23 February 1956, which yields the spectrum of relativistic protons $\mathbf{P}^{-7}$.[13] This is indicated by (1) in the Figure, and for comparison the energy spectrum of normal cosmic rays is also shown. The late effect of the solar flare on 23 February 1956 gives further information about the spectrum for the non-relativistic

can be obtained from various phenomena, such as polar cap blackouts, type III absorptions of VHF cosmic radio waves and high altitude observations of cosmic radiation by satellite. Curve (3) is based on results of Winckler *et al.*[6] from their balloon observations of solar cosmic rays at Churchill. Most other values mentioned are ranging between curve (2) and (3).

Below 1 MeV, Chamberlain's result (curve 4) from the observations of Doppler shift of the auroral spectral lines is available.[16] Direct

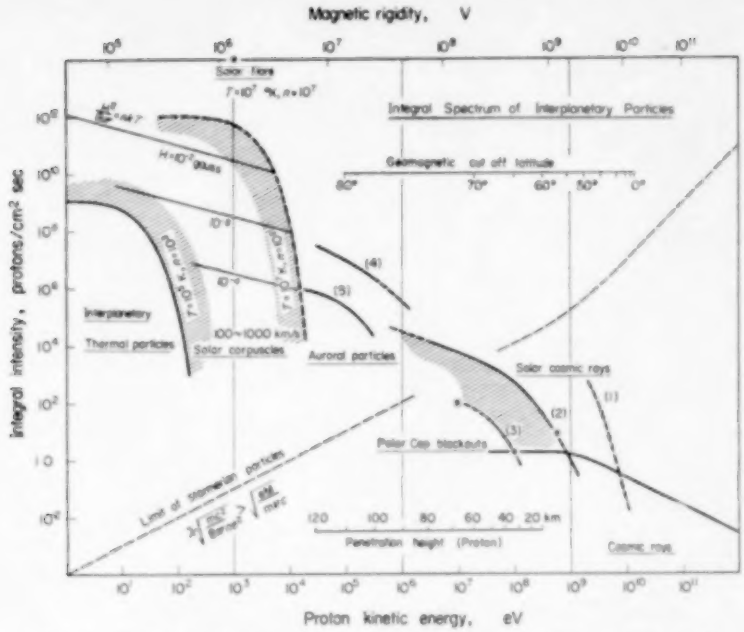


Fig. 4. The integral energy spectrum of solar particles associated with intense solar flares.

rocket observations of auroral particles have been made;[17, 18] however, there seems to remain still some ambiguities. Corpuscular streams consist mainly of rather low energy particles and travel interplanetary space with a speed of the order of 1000 km/sec. Chapman[19] has given an estimate of the density of a neutral ionized gas stream—100 protons and electrons/cm³—sufficient to cause a moderate magnetic storm resulting in a velocity of 2000 km/sec. On the other hand, Biermann[20] has pointed out that the motion of comet tails indicates the existence of corpuscular clouds streaming away from the sun with velocities of 500–1500 km/sec. These observations lead to a value of the particle flux near the earth of some 10^{10} cm^{-2} sec^{-1} under quiet conditions, but up to 10^{12} cm^{-2} sec^{-1} in magnetic storms. These values are in good agreement with the estimates from the theory of formation of geomagnetic cavities surrounded by corpuscles during magnetic storms.[15] The minimum flux of solar corpuscles may be indicated by the density of the interplanetary gas. In his theory of the solar outer corona, Chapman[21] showed that the temperature of the gas is about 2×10^5 °K (30 eV) near the earth, and the density may be of the order of 600/cm³ according to the observations of zodiacal lights.[22] These values yield a flux of particles of about 10^9 cm^{-2} sec^{-1}.

Little is known about the spectrum of the solar corpuscles of lower energies. However, there is an evidence that they might be in a state of kinetic equilibrium. This condition can be studied by examining a characteristic collision time of particles. According to Spitzer[23], a measure of the time required for the distribution of kinetic energies of particles to approach the Maxwellian distribution is given by

$$\tau_c \simeq \frac{11{\cdot}4\, T^{3/2}}{n \ln \Lambda} \qquad (3)$$

for proton gas, where T is in degrees K and $\ln \Lambda \simeq 20$ at most conditions in the solar atmosphere. Inserting numerical values of

$T \sim 10^6$–10^7 °K and $n \sim 10^7/\text{cm}^3$ in the corona, it can be shown that the Maxwellian distribution is attained within a few hours. For electron gas this characteristic time is less than that for protons by a factor of 1/43. Therefore one may conclude that solar corpuscles in the lower energy range (say less than 1 keV) are in kinetic equilibrium and hence Maxwellian distribution is attained.*

This result has an interesting consequence. Since the corpuscular cloud undergoes an expansion, the particles in the cloud suffer adiabatic cooling. If the expansion is slow compared to τ_c, then

$$T \propto n^{\gamma-1} = n^{2/3} \tag{4}$$

Therefore the mean kinetic energy of particles would decrease from 1 keV to 10 eV as the particle density decrease from $10^7/\text{cm}^3$ to $10^4/\text{cm}^3$. It is to be noted that, though the speed of the bulk of corpuscular cloud is about 1000 km/sec, energies of individual particles or their thermal velocities become considerably low.

Although most low-energy particles less than 1 keV undergo adiabatic cooling, high energy particles would not suffer this because of their high velocities and low particle density, though they might be accelerated in the corona by Fermi mechanism. Recently Parker and Tidman[24] have discussed the generation of suprathermal particles in a violently agitated plasma carrying magnetic fields. According to their analysis, in order that thermal particles can be raised to suprathermal ones, the characteristic time for the Fermi process, τ_F, must be shorter than that of the Coulomb interaction time, τ_c, i.e.

$$\tau_F = \frac{Lv}{V^2} < \tau_c = \frac{\lambda(v)}{v} = \frac{m^3v^5}{32\pi e^4 nkT \ln\Lambda} \tag{5}$$

where L is the mean free path for the collision of a particle within magnetic clouds and V is the velocity of hydromagnetic waves which are usually comparable to the thermal velocity. Then the threshold energy E^* is given, putting $L \sim 10^4$ km and $n \sim 10^7/\text{cm}^3$,

$$E^* > \left[\frac{8\pi}{3} e^4 nL \ln\Lambda\right]^{1/2} \sim 0{\cdot}3 \text{ keV} \tag{6}$$

Therefore most thermal particles with energy exceeding above 1 keV would undergo Fermi acceleration to form a high-velocity non-Maxwellian suprathermal tail. For lower energy particles, however, the slowing down by Coulomb interaction is so effective that they remain in the Maxwellian distribution.

The energy spectrum of such suprathermal particles can be computed theoretically assuming that particles will undergo N_0 collisions in average with moving magnetic inhomogeneities before being expelled from the cloud. According to Parker[12], if the Fermi acceleration is operative, the integral energy spectrum $\mathbf{J}(>E)$ near the relativistic range is given by

$$\mathbf{J}(>E) = 2^\gamma J_0 \left[1 + \frac{(E+m_0c^2)^2}{m_0^2c^4}\right]^{-\gamma} \tag{7}$$

where $\gamma = c^2/(4N_0V^2)$ and m_0c^2 is the rest energy of a proton. For extreme relativistic particles, $\mathbf{J} \sim E^{-2\gamma}$. As has been shown, the observed spectrum at the relativistic range is about $\mathbf{E}^{-5}$, yielding the value $\gamma \simeq 2{\cdot}5$.

In Fig. 4, the vertical penetration depths and

---

* If the particle density n/cm^3 is given with mean random velocity v and the translational velocity perpendicular upon a unit fixed surface with V, then the integrated flux is

$$\mathbf{J}(>E) \simeq \int_E^\infty nV\,dE$$

when $v \ll V$. Assuming a Maxwellian distribution, then

$$\mathbf{J}(>E) = n_0 V\left[1 - \frac{2}{\sqrt{\pi}}\int_0^x \exp(-x^2)\,dx + \frac{2}{\sqrt{\pi}}\, x\exp(-x^2)\right]_{x=\sqrt{\left(\frac{E}{kT_0}\right)}}$$

where n_0 is the total number of particles in a unit volume and $v = \sqrt{\left(\frac{2kT_0}{m}\right)}$. For the case when $v \gg V$,

$$\mathbf{J}(>E) = \frac{n_0 v}{\sqrt{\pi}}\left[(1+x^2)\exp(-x^2)\right]_{x=\sqrt{\left(\frac{E}{kT_0}\right)}}$$

magnetic cut-offs are also indicated in order to have some idea of the range-vs.-energy characteristics of typical particles penetrating into the lower ionosphere. The penetration depths are due to Bailey[1] and the cut-off magnetic rigidities for the indicated geomagnetic latitudes are those for an ideal geomagnetic dipole field given by the equation

$$P = 14{\cdot}7 \cos^4 \Phi, \text{(in BV)} \qquad (8)$$

Another very interesting criterion for charged particles is the one which gives the condition of whether the particles entering the geomagnetic field behave like a single particle or a conducting plasma. This criterion has been investigated by Ferraro[25], who showed that below a certain critical density of particles the electro-magnetic interaction between particles becomes negligible and their motion will, then, be very nearly Störmerian. The condition can approximately be obtained by equating the shielding depth of a plasma λ_D and the impact parameter of a particle in the geomagnetic field (Störmer's unit length), i.e.

$$\lambda_D = \sqrt{\left(\frac{mc^2}{8\pi ne^2}\right)} \geqslant \sqrt{\left(\frac{eM}{mvc}\right)} \qquad (9)$$

where n is the particle density and M is the geomagnetic moment. Then, the critical flux of particles is given by

$$\mathbf{J}_c = \frac{1}{4} n_c v = \frac{c}{32\pi eM}\left(\frac{P}{300}\right)^2 \qquad (10)$$

where P is the rigidity in volts. As is indicated by a dotted line, this limit crosses over the spectrum of solar particles at a range of about 10–100 MeV. Therefore above this energy range, solar particles can be treated as independent particles and the scale of the geomagnetic cut-offs shown in the figure is applicable, while lower energy particles below this limit must be regarded as a conducting gas.

4. CONCLUSIONS

Important evidence indicating the existence of energetic solar particles associated with solar flares is given. It has now become possible to investigate the whole spectrum of solar particles, and many individual observations such as cosmic rays, solar radio emissions, cosmic radio waves, aurorae, ionosphere and geomagnetism can be co-ordinated. Though the present aspects about solar corpuscular emissions and their effects in the earth's upper atmosphere are consistent with the results obtained in various fields, the investigation is still crude and further observational and theoretical analyses are desired.

In conclusion, the authors wish to express their thanks to the members of the Ionosphere Research Committee for valuable discussions.

REFERENCES

1. D. K. Bailey, *Proc. Inst. Radio Engrs., N.Y.* **47,** 255 (1959).
2. Y. Hakura, Y. Takenoshita and T. Otsuki, *Rep. Ionos. Res., Japan* **12,** 459 (1958).
3. T. Obayashi and Y. Hakura, *J. Atmos. Terr. Phys.*, in press; and *J. Radio Res. Lab. Japan,* **7,** 17 (1960).
4. G. C. Reid and C. Collins, *J. Atmos. Terr. Phys.* **14,** 63 (1959).
5. H. Leinbach and G. C. Reid, VHF radio wave absorption in northern latitude and solar particle emissions. Private communication, May (1959).
6. J. R. Winckler, R. Arnoldy, R. Hoffman, L. Peterson and K. A. Anderson, *Bull. Amer. Phys. Soc.* **4,** 238 (1959).
7. P. Rothwell and C. McIlwain, *Satellite Observation of Solar Cosmic Rays, Iowa Univ. Rep.* SUI-59-12 (1959).
8. K. Sinno and Y. Hakura, *Rep. Ionos. Res., Japan* **12,** 285 (1958).
9. A. Boischot and J. F. Denisse, *C.R. Acad. Sci., Paris* **245,** 2194 (1957).
10. Y. Hakura and T. Goh, *J. Radio Res. Lab., Japan* **6,** No. 28 (1959).
11. S. Hayakawa and K. Kitao, *Prog. Theor. Phys., Osaka* **16,** 139 (1956).
12. E. N. Parker, *Phys. Rev.* **107,** 830 (1957).
13. P. Meyer, E. N. Parker and J. A. Simpson, *Phys. Rev.* **108,** 1563 (1957).
14. S. Chapman and V. C. A. Ferraro, *Terr. Magn. Atmos. Elect.* **37,** 147 (1932).
15. T. Obayashi, *Rep. Ionos. Res., Japan* **12,** 301 (1958).
16. J. W. Chamberlain, *Astrophys. J.* **120,** 360 (1954).
17. J. A. Van Allen, *IGY Rocket Rep. Ser.* No. 1, 159 (1958).
18. L. H. Meredith, L. R. Davis, J. P. Heppner and O. E. Berg, *IGY Rocket Rep. Ser.* No. 1, 169 (1958).

19. S. CHAPMAN, *J. Geophys. Res.* **55,** 361 (1950).
20. L. BIERMANN, *Observatory* **77,** 109 (1957).
21. S. CHAPMAN, *Smithson. Contr. Astrophys.* **2,** No. 1 (1957).
22. H. SIEDENTOPF, A. BEHR and H. ELSASSER, *Nature, Lond.* **171,** 1066 (1953).
23. L. SPITZER, *Physics of Fully Ionized Gases.* Interscience Publishers, New York (1956).
24. E. N. PARKER and D. A. TIDMAN, *Phys. Rev.* **111,** 1026 (1958).
25. V. C. A. FERRARO, *J. Geophys. Res.* **57,** 15 (1952).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 70-78. Printed in Great Britain.

RESEARCH NOTES

Measurement of the Net Electric Charge on the Sun by means of the Artificial Planet "Pioneer V"

(*Received* 30 *March* 1960)

It is generally assumed, without strict proof, that bodies like the sun do not carry appreciable net electric charges. On the other hand it can be shown that the *hypothesis* that such a body of mass M g carries a negative charge $-Q$ e.s.u., where

$$Q=\beta G^{\frac{1}{2}}M, \tag{1}$$

$\beta \sim 0{\cdot}03$ and G is the constant of gravitation, serves to derive the known orders of magnitude of the following six different astronomical phenomena and the directions relating to three of them:

(1) the present magnetization of the earth;

(2) the outer Van Allen belt of radiation;

(3) the maximum energy, about 5×10^{15} eV, of a primary cosmic ray particle;

(4) the sun's north polar magnetic field;

(5) the approximate equality of the ratio P/U for the sun to the mean value of this ratio for Blackett's group of five magnetic stars, where P is the magnetic moment and U is the angular momentum of a star;

(6) the magnetic field in that region of Jupiter which emits strong, elliptically-polarized radio bursts on frequencies about 20 Mc/s.

It is therefore highly desirable to use the opportunity which is now provided by the magnetometric measurements made in the artificial planet Pioneer V (PV) to determine the value of the net charge on the sun. Its feasibility may be shown as follows.

In a non-rotating frame of reference attached to PV the sun's charge $-Q$ moves with a velocity $-v$ where v is the velocity of PV at a given point R in its orbit about the sun. Then, on account of Ampère's law or the Lorentz transformation, there will exist in this frame a magnetic field H_r which is perpendicular to the orbital plane, is directed from north to south* and has near the magnetometer the intensity H_r where

$$H_r=(Q/c)\,v\sin\phi/r^2,$$

r is the distance of R from the sun and ϕ is the angle between the radius vector $\mathbf{r}$ and the velocity vector $\mathbf{v}$. Since the theory of planetary orbits shows that the areal velocity $\frac{1}{2}vr\sin\phi$ of a planet about the sun is a constant $\frac{1}{2}h$, it follows that

$$H_r=A/r^3 \tag{2}$$

where $A=hQ/c$ is a constant.

It is remarkable that the formula (2) also represents the field in the orbital plane set up by a dipole of magnetic moment A situated at the sun's centre with its axis normal to this plane.

These results show that if Q is large enough its value could be determined from the measurements of H_r made by the transported magnetometer† and that at the same time the observed dependence of H_r on r could be used as a test of the theory.

Since (as reported) the orbit of PV will lie between the orbits of the earth and of Venus the values of H_r will all be measurable if $H_r>10^{-5}$ Gauss when $r=r_E$ the radius of the earth's orbit. Hence, by equation (2), Q will be detectable when

$$Q>10^{-5}r_E^{\,3}c/h. \tag{3}$$

Since also the value of h for PV is of the same order of magnitude as that for the earth, which is $4{\cdot}5\times10^{19}$, the smallest detectable charge on the sun is therefore about $2{\cdot}2\times10^{25}$ e.s.u.

Under the hypothesis (1) the charge on the sun is $-Q_S$ where $Q_S\sim1{\cdot}5\times10^{28}$ e.s.u. Hence

* Since, presumably, PV moves around the sun in the same direction as that of the earth.

† Provided that any screening effect of the metal envelope of PV is negligible or else circumvented in some convenient way.

that charge would be easily measurable if our hypothesis were true. The corresponding minimum and maximum values of H_r would then be about 7 and 14 mG respectively.

As shown elsewhere,[1] this value of Q_s serves to derive the orders of magnitude and relevant directions of the phenomena (1)–(5) mentioned above.

When hypothesis (1) is also applied to the planets Jupiter and Saturn, which resemble the sun chemically and in being subject to an occasional eruption, the intensities of the polar magnetic fields set up by their rotations are respectively about 5 and 1·7 G.

The value 5 G for Jupiter agrees in order of magnitude with the estimated values >4 G and about 7 G derived by Franklin and Burke[2] and by Smith and Carr[3] respectively from observations of the elliptically-polarised radio noise emitted from Jupiter. This agreement has already been mentioned above under (6).

Our estimate of 1·7 Gauss for the polar field of Saturn suggests (by analogy with Jupiter) that any non-thermal noise which Saturn might emit would be much stronger at frequencies near 5 Mc/s than at those near 20 Mc/s which have recently been used to detect noise from Saturn.[3] If this suggestion could be verified it would constitute additional evidence in support of the hypothesis (1).

References

1. V. A. Bailey, *Nature, Lond.* **186**, 508 (1960); *J. Atmos. Terr. Phys.*, in press; *J. Roy. Soc. N.S.W.*, in press.
2. K. L. Franklin and B. F. Burke, *J. Geophys. Res.* **63**, 807 (1958).
3. A. G. Smith and T. D. Carr, *Astroph. J.* **130**, 641 (1959).

V. A. Bailey

Research Professor,
University of Sydney

Radar Echoes from Venus and a New Determination of the Solar Parallax

(*Received* 1 *April* 1960)

An earlier communication[1] described an experiment in which a radar transmission was directed at the planet Venus in 1958 near its time of closest approach to the earth, and the presence of returning pulses was subsequently detected in the recorded data. In this experiment it was possible to make unambiguous measurements of round-trip propagation times on two different days. The present note reports the derivation of new values of solar parallax and astronomical unit from these two early observations, and also describes an attempt to repeat the experiment.

1. SOLAR PARALLAX CALCULATIONS

In the earlier experiment the propagation time was measured on 10 and 12 February 1958, and the values obtained were reported to be 295·5065 sec and 302·9842 sec, respectively.[1] For 10 February, the travel time just given was that elapsing between the transmission of the first pulse from the Millstone Hill radar site at 14:21:05 EST and the later time of its reception back at the site. For 12 February, the corresponding time of transmission was 12:00:55 EST. The Millstone site geodetic latitude is 42°37′3″ N and the longitude is 71°29′32″ W.

From the above parameters and standard values for the velocity of light, equatorial radius and shape of the earth, and radius of Venus, we have derived values for the size of the astronomical unit (mean radius of the earth's orbit) and solar parallax (angle subtended at one astronomical unit distance by the earth's equatorial radius). We assume that the radar reflection takes place from a point closer to the earth than the center of Venus by an amount equal to the adopted Venus radius; simple corrections to the results derived below will doubtless become necessary when more is known about the reflecting characteristics of Venus.

In our method of computation we find for each observation the distance, in astronomical units, existing between the earth and Venus centers at the time that reflection of the first pulse of our 5-min long transmission took place

(taking into account the significant amount of relative earth–Venus motion that occurred during the propagation of the signal). The total distance traversed by a fictitious pulse, emitted from the center of the earth at the specified starting time and reflected back to the center of the earth from the center of Venus, is then double the one-way distance thus found. By correcting the observed travel time for the earth and Venus effective surface-to-center distances, a travel time for the fictitious pulse is obtained. The size of the astronomical unit is then obtained by dividing the product of the corrected travel time and the velocity of light by the round-trip distance, in astronomical units, found for the fictitious pulse. An approximation is involved in assuming that the total travel distance is twice the separation at the instant of arrival at Venus of the pulse. This, however, is believed to produce second-order errors which are negligible, because earth and Venus orbital velocities are so small a fraction of the velocity of light.

The planetary center-to-center distances (in astronomical units) at the time of reflection have been found from four-point Lagrangian interpolation on daily earth–Venus center-to-center distances existing at 0 hr Ephemeris Time (which in 1958 was 5 hr 33 sec later than EST),[2] together with a linear equation accounting for the relative planetary motion during the flight of the pulse. The daily distances in turn were obtained after Lagrangian interpolation on standard ephemerides.[3] These tables give values for the radius vectors, and latitudes and longitudes referred to the mean equinox and ecliptic of date, taken at 4-day intervals for the sun as seen from the earth and for Venus as seen from the sun (all references being the centers of the respective bodies). The daily earth-Venus distances are related through straightforward trigonometric calculations to the daily values of the above polar coordinates. We have found by this procedure that the round-trip distance that would have been traversed by the fictitious pulse on 10 February was 0·5928142 astronomical units, while for 12 February it was 0·6078308 astronomical units.

In making the effective surface-to-center correction for the earth, it can safely be assumed that Venus is infinitely distant. The distance correction is thus equal to the sum of the length of the component of the Millstone-to-earth-center vector that lies in the direction of the earth–Venus center-to-center vector at the instant that the radar transmission commenced, plus the length when the first reflected pulse was received. The component is of length

$$C = R\,(\cos d \cos \lambda \cos \mathrm{LHA} + \sin d \sin \lambda)$$

where R is the Millstone-to-earth-center distance, d is the declination of Venus, λ is the geocentric latitude of the Millstone site, and LHA is the local hour angle of Venus at Millstone at the appropriate instant. The declination and local hour angle are found by Lagrangian interpolation in standard astronomical tables,[4] and the geocentric latitude of the site has been determined using the International Ellipsoid of Ref. (4). We find for Millstone: $R = 6368{\cdot}59$ km, $\lambda = 42^\circ 25' 30''$ N. For the observation of 10 February, the total earth surface-to-center correction is determined to be 6300 km, while for 12 February it is 13,038 km.

The total correction to be added to the observed travel time is given by adding to the surface-to-center correction for the earth twice the assumed Venus radius, and dividing the sum by the velocity of light. The radius of Venus is taken to be 6123 km, a recent value due to Kuiper.[5] The value adopted for the velocity of light is 299,792·5 km/sec.[6] The total corrections thus found for 10 February and 12 February are, respectively, 0·0514 sec and 0·0626 sec.

After applying the correction to the observed travel time, multiplying the sum by the above velocity of light* and equating the result to the round-trip distance in astronomical units found earlier for the fictitious pulse, we find that the observation of 10 February implies that one astronomical unit $= 149{\cdot}4668 \times 10^6$ km. For 12 February, the result is: one astronomical unit $= 149{\cdot}4678 \times 10^6$ km. The reason for the

* The correction in velocity caused by an upper limit of 10^3 electrons/cm^3 in the interplanetary void is less than 1×10^{-6} at this frequency and may be neglected.

discrepancy between the results of the two days, which is equivalent to about 0·002 sec in travel time, is not apparent. Using for the equatorial radius of the earth the value of 6378·388 km adopted by the International Astronomical Union,[4] we conclude that the average value of the solar parallax as measured on the two days is $8''\cdot8021 \pm 0''\cdot00006$.

2. OBSERVATIONS OF AUGUST–SEPTEMBER 1959

For a period of 4 weeks surrounding the inferior conjunction of Venus in early September 1959, attempts were made to repeat the radar echo measurements of February 1958. In these attempts two approaches were followed. The first, comprising a total of sixty-six valid 5-min runs, obtained coherent, digitalized recordings of the returned signal for subsequent processing by a computer. This method was almost identical to that used in February 1958.[1] The second, with a total of 117 5-min runs, involved initial analogue processing in a matched filter and squarer followed by digitalization and incoherent integration in real time by the CG 24 computer at the radar site. Although lacking in flexibility and sophistication, the latter approach was included because of its substantially lower processing expense and its ability to display processed results immediately.

Table 1.

Operating radio wavelength	$\lambda = 0{\cdot}681$ m
Peak transmitted power	$P_t = 500$ kW
Pulse duration	$t = 2$ msec
Pulse repetition rate	PRF $= 50$ sec^{-1}
Overall system noise temperature	$T = 185°$K
Antenna gain (over isotropy)	$G = 37$ dB
Transmitted polarization	right circular
Received polarization	left circular
System losses	$L = -2$ dB

Typical radar parameters used during this second series of attempts are shown in Table 1. In only one respect do they differ materially from those of February 1958: a higher pulse repetition rate (at no sacrifice in peak power) was used in the hope of improving signal detectability. A lower sideband up-converter using a variable reactance diode was substituted for the maser used previously. This device did not yield quite so low a noise temperature as the maser but was gratifyingly stable and reliable in its operation. A hoped-for increase in transmitted power was not realized in time for this experiment.

The total number of valid, independent runs amounted to 138, of which 45 were observed simultaneously by the two approaches. It was verified that correlation curves produced by the two methods were in excellent agreement. Of the total number of runs, only one displayed a peak sufficiently above noise to be statistically significant. This peak, 4·2 standard deviations high, was observed at 1916 UT on 2 September 1959 using both processing techniques and differed by only 1·1 msec* from a prediction based on the findings of February 1958. The probability that noise alone would produce a peak of this magnitude or higher in any range box over the total runs observed is about 2 per cent. In order to study this event in more detail, the 5-min sequence was broken into four parts, each of which were separately analyzed. The first section yielded a very strong return (which, however, was not due to any small number of strong signals) at precisely the predicted range and frequency. The second and third sections contained no important peaks, while the fourth showed a peak of 4·4 standard deviations centered 1750 c/s below the predicted doppler shift at a range 2 msec greater than predicted. Since there is no explanation for so large a frequency deviation which can be considered acceptable at the present time, this event has been cast out. We must, therefore, conclude that none of the individual runs show strong evidence of Venus echoes.

In the hope that still further integration might produce a recognizable signal, the results from

* In order to insure maximum detectability, this run employed a simple repetitive transmission rather than a coded sequence of pulses. However, as a result, the measurement of round-trip travel time is uncertain to within a multiple of the interpulse period.

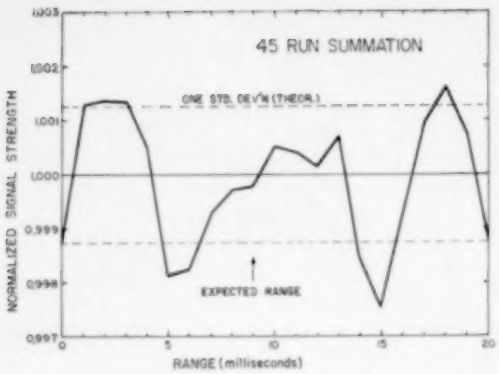


Fig. 1.

forty-five of the digitally-recorded runs were combined in register and are shown in Fig. 1. No statistically significant peak is in evidence. If an upper limit is assigned to a signal which could not be detected, the corresponding (steady) incoming signal power and, therefore, an upper limit for the target radar cross-section may be calculated. Let S equal the signal voltage output of the integrator expressed in standard deviations of the accompanying noise. The radar cross-section of Venus (in units of its projected area) which would yield such a signal may then be calculated from

$$\sigma' = 32\left(\frac{S}{\sqrt{n}}\right)\frac{\pi^3 R^4 kT}{P_t t G^2 \lambda^2 V^2 L}$$

where n equals the total number of samples which have been integrated, k is Boltzmann's constant, R is the distance to the planet and V the Venus radius. The remaining quantities are listed in Table 1. For the grand summation, n equals $6{\cdot}3 \times 10^5$ and we find:

$$\sigma' = 0{\cdot}026S$$

For example, $S=2$ (a result which noise has 0·37 probability of producing at least once in the ten channels here) gives a corresponding upper limit on the radar cross-section of 0·052. While this result is some 13 dB smaller than the tentative findings of February 1958, it is quite consistent with the measurements made at Jodrell Bank, England,[7] during the present conjunction, and is approximately what would be expected if Venus exhibited scattering properties similar to our Moon. It is certainly too early to attempt an explanation for the reduced detectability of the planet during the recent observations as compared to February 1958. However, it should be remembered that even on that occasion evidence of radar returns was found in only two of the four valid runs.

Acknowledgements—Acknowledgement should be made of the efforts of G. Westrom and his associates at Aeronutronic Systems, Inc. in the independent checking of the solar parallax calculations. The direct participation of H. Briscoe, P. Fleck, L. Kraft and W. Smith in the experiment was invaluable; without their aid and interested involvement such a complex undertaking could not have been carried out. Supporting equipment and computer routines were developed by J. Craig, T. Goblick, J. Kolb, A. Parsons and B. Tague.

G. PETTENGILL
R. PRICE

Lincoln Laboratory,
Massachusetts Institute of Technology,
Lexington 73,
Massachusetts

References

1. PRICE *et al.*, *Science* **129,** 751 (1959).
2. U.S. Naval Observatory, *The American Ephemeris and Nautical Almanac for the Year 1960.* U.S Government Printing Office, Washington, D.C. (1958).
3. *Astronomical Papers of the American Ephemeris* Vol. 14, Solar Coordinates 1800–2000, and Vol. 15, part 3, Coordinates of Venus 1800–2000. U.S. Government Printing Office, Washington, D.C. (1953 and 1955, respectively).
4. U.S. Naval Observatory, *The American Ephemeris and Nautical Almanac for the Year 1958.* U.S. Government Printing Office, Washington, D.C. (1956).
5. G. KUIPER, *The Atmospheres of the Earth and Planets.* University of Chicago Press (1954).
6. Value adopted by the International Union of Geodesy and Geophysics; also the value measured by K. D. FROOME, *Nature, Lond.* **181,** 258 (1958).
7. J. V. EVANS and G. N. TAYLOR, *Nature, Lond.* **184,** 1358 (1959).

On Emission λ 10830 Å in Aurorae

(*Received* 1 *July* 1960)

A considerable enhancement of the intensity of the *Q*-branch (5, 2) OH (λ 10830 Å) relative to the intensity of the *R*-branch of the same band was observed in Zvenigorod during the great aurora of 10–11 February 1958[1]. However, this single observation has not been confirmed up to now.

Since January 1960 observations of spectra of aurorae have been carried on in a wide spectral region including the infra-red region in Loparskaya ($\phi = 68°\ 38'$; $\lambda = 33°\ 22'$; $\Phi = 63°\ 37'$; $\Lambda = 126°\ 42'$). 1960 was rather poor in aurorae. The aurorae were weak and lasted for a short time; thus spectra of the aurorae and of the night sky were found combined. But on March 31–1 and on April 1–2, 1960 one could observe bright forms: arcs, violet rays and red spots during the whole night (i.e. during the morning and night twilight) mostly in the south. The spectrograph was directed to the south and was guided by the bright forms in the vertical. All the emissions of the aurorae, in particular atomic lines, were notably enhanced at the spectra. The spectra were very similar to that obtained in Zvenigorod on February 10–11, 1958. Besides the enhancement of λ 10830 Å, a very strong line ^2D–^2P NI λ 10400 Å was present.

The average relation between the *Q*- and the *R*-branches of the band (5, 2) OH in an ordinary sky, as follows from the observations at the Burakan Observatory ($\phi = 40°\ 21'$ N), is equal to 1·35 and never exceeds 1·45[2]. Thus, any relation exceeding 1·5 was considered to be due to the helium emission.

A preliminary treatment of the results obtained in 1960, showed that the relation of the intensities of the *Q*- and *R*-branches of the band (5, 2) OH is approximately equal to unity in the night sky in the north while in the aurorae one can frequently observe a relation exceeding two. In the case observed on March 31–1 and April 1–2, 1960, this relation was equal to 3·5 and 3·2 respectively. In the last two cases there were sunlit aurorae as in the case of the observations in Zvenigorod. In the other cases the aurorae were partially sunlit.

In conclusion, I express my thanks to Prof. Krassovsky for a discussion of the results.

N. I. FEDOROVA

Institute of Physics of the Atmosphere,
Academy of Sciences of U.S.S.R.,
Bol. Gruzinskaja 10,
Moscow

References

1. A. V. MIRONOV, V. S. PROKUDINA, N. N. SHEFOV. *Spectral, Electrophotometrical and Radar Researches of Aurorae and Airglow*, No. 1, 20 (1959).
2. N. I. FEDOROVA, *Izvest. Akad. Nauk. SSSR, ser. Geofis.*, No. 6, 836 (1959).

On the Nature of Helium Emission λ 10830 Å in Aurorae

(*Received* 1 *July* 1960)

During the period when extremely strong aurora occurred on February 10–11 1958 in Zvenigorod ($\phi = 55°\ 43'$; $\lambda = 36°\ 51'$; $\Phi = 51°\ 03'$; $\Lambda = 120°\ 16'$) there was found a general increase of emission at the *Q*-branch of the hydroxyl band (5, 2) as compared to the *R*-branch[1]. It was provisionally assigned to the helium emission: $2^3P - 2^3S$. At present, similar emission have been more than once registered during the aurorae in Loparskaya ($\phi = 68°\ 38'$; $\lambda = 33°\ 22'$; $\Phi = 63°\ 37'$; $\Lambda = 126°\ 42'$)[2,3]. In this period when there was an increased emission of λ 10830 Å there appeared red aurorae lighted by the Sun in Zvenigorod as well as in Loparskaya.

Now, since the additional emission of λ 10830 Å had been repeatedly observed, and, consequently, is not a result of experimental error it would be interesting to try to understand the nature of this emission. It seems probable that it is due to the fluorescence of helium atoms in the metastable state 2^3S from the Solar emission. It is most likely that such metastable atoms originate as a result of the bombardment of neutral atoms of helium with electrons possessing energies close to the energy of the excited state of He 2^3S (~20–25 eV). These electrons arise in the zone of appearance

of red aurorae at a height of 200–300 km by ionization of the atmospheric atoms and molecules due to electrons with an energy of about 10 keV as had already been discovered directly[3, 4, 5].

The number of light quantas with λ 10830 Å, arising in a column with a base of one cm^2 and denoted by N can be represented:

$$N = H \cdot n_{He} \cdot n_{e^*} \cdot V \cdot \sigma \cdot \tau \cdot \nu$$

where H is the height of the homogeneous helium atmosphere, n_{He} is the concentration of neutral helium atoms, n_{e^*} is the concentration of electrons with an energy of about 25 eV, V is the velocity of these electrons, σ is the effective cross section for helium excited in the state 2^3S by these electrons, τ is the average life-time of metastable He2^3S and ν is the number of quanta of the Solar emission scattered by helium in this state per sec. Generally these values can be calculated with sufficient accuracy. However, here we confine ourselves only to approximate estimates allowing variations in detail, to show the probability of an assumed mechanism for the emission of He 2^3P-2^3S. If one accepts that at a height of 200–300 km H $\sim 10^7$ cm, $n_{He} \sim 10^6$ cm^{-3} $n_{e^*} \sim 10^2$ cm^{-3} $V \sim 3 \times 10^8$ cm sec^{-1} $\sigma \sim 10^{-15}$ cm^2, $\tau \sim 10^3$ sec and $\nu \sim 10$ sec^{-1}, one obtains $N \sim 3 \times 10^9$ cm^{-2}. This value agrees well with that observed during the aurora on February 10–11, 1958 in Zvenigorod.

The energy flux carried by electrons which excite helium (~ 25 eV) will be of the order of 1 erg cm^{-2} sec^{-1}. There is nothing incredible in this, since the electron flux creating them with an energy of 10 keV can penetrate during lower magnetic disturbances and is generally more powerful even under calm geomagnetic conditions[3, 4, 5].

Thus, there is a fascinating perspective for investigating some characteristics of the upper atmosphere by the helium emission 2^3P-2^3P. There are all reasons to expect that in the zone of its appearance we will be able to estimate the concentration of helium as well as the concentration of electrons having from 25 eV to tens of keV of energy.

The author expresses his thanks to Prof. V. I. Krassovsky for his attention and advice while carrying out this work.

N. N. SHEFOV

Institute of Physics of the Atmosphere, Academy of Sciences of U.S.S.R., Bol. Gruzinskaja 10, Moscow.

References

1. A. V. MIRONOV, V. S. PROKUDINA, N. N. SHEFOV, *Spectral, Electrophotometrical and Radar Researches of Aurorae and Airglow*, No. 1, 20 (1959)
2. N. I. FEDOROVA, *Planet Space Sci.* preceding note.
3. N. I. KRASSOVSKY, IU. M. KUSHNIR, G. A. BORDOVSKY, G. F. ZAKHAROV, E. M. SVETLITSKY, *Fifth General Assembly of CSAGI, Moscow, Aug.* 1958. *Iskusstvennye Sputniki Zembi No.* 2, 59 (1958).
4. V. I. KRASSOVSKY, I. S. SHKLOVSKY, G. I. GALPERIN, E. M. SVETLITSKY, *Izvest. Akad. Nauk ser. Geofis.*, No. 8, 1157 (1959).
5. V. I. KRASSOVSKY, I. S. SHKLOVSKY, G. I. GALPERIN, E. M. SVETLITSKY, *Dokl. Akad. Nauk SSSR*, **127**, No. 1, 78 (1959).

Origin of the *D*-Region

(*Received* 4 *May* 1960)

The recent results reported by Jursa, Tanaka and LeBlanc,[1] in their attempt to detect NO in the *D*-region of the ionosphere, appear to cast some doubt as to the role of this constituent in the formation of the *D*-region. They conclude that the NO concentration in the altitude interval 63 to 87 km cannot exceed $10^8/\text{cm}^3$. Whereas, Nicolet[2] has suggested that concentrations of $10^8/\text{cm}^3$ near 80 km or 10^9 near 65 km appear to account for the ionization rates required in the *D*-region. It thus seems evident that although photoionization of nitric oxide by solar Lyman-α radiation could account for the formation of the *D*-region, all experiments to date have failed to detect the requisite amount of NO. In view of this, it may be appropriate to inquire whether some other constituent or mechanism, not heretofore suggested, may indeed be responsible for the formation of the *D*-region of the ionosphere.

The following suggestion is offered as an alternate hypothesis of a possible mechanism in the formation of the *D*-region. It is well known that vibrationally excited O_2 in the ground electronic state is metastable. Furthemore, in shock-tube studies by Blackman[3] the measured vibrational relaxation time of O_2 was 54 μ sec at 800°K and 1 atm pressure. This latter result indicates a relatively low probability of collisional deactivation of vibrationally excited O_2. Therefore, the lifetime of these excited molecules at densities corresponding to the *D*-region (about $10^{15}/cm^3$) could be as much as a few tenths of a second. We now consider an O_2 molecule in the ground electronic state with, say, 11 vibrational quanta. In accordance with the potential energy curves of $O_2(^3\Sigma_g^-)$ and $O_2^+(^2\Pi_g)$ the following Franck-Condon transition is certainly energetically possible

$$O_2(^3\Sigma_g^-)_{v''=11} + h\nu \longrightarrow O_2^+(^2\Pi_g)_{v'=0} + e^- \qquad (1)$$

where the ionization potential of O_2 is about 12·1 eV[4] and the excitation energy of $O_2(^3\Sigma_g^-)_{v''=11}$ about 1·9 eV. From this we infer that $h\nu$ must be about 10·2 eV, the wavelength corresponding to this energy being 1216 Å, which is just that of Lyman-α. We also note that the above transition holds for O_2 molecules with more than 11 vibrational quanta, i.e. for $v'' > 11$. For transitions arising from the latter, maximum probability would occur for those ending in $O_2^+(^2\Pi_g)$ with $v' > 0$ in accordance with the Franck–Condon principle. Optically allowed transitions have been observed in the Schumann–Runge system in absorption for vibrationally excited O_2,[5, 6, 7] which will certainly compete with the above ionizing transitions.

We now ask what mechanisms are possible which could populate the *D*-region with vibrationally excited O_2. Certainly thermal excitation is not involved since the temperature of this region is about 200°K. One possible mechanism is as follows. All transitions in absorption in the Schumann–Runge system which excite O_2 into the $^3\Sigma_u^-(v'=0,1)$ state will, in the subsequent downward transitions (emission), tend to populate the $O_2(^3\Sigma_g^-)$ molecules with $v''=10$ to 17. This is a consequence of the large difference in internuclear separation of the two electronic states and the Franck–Condon principle.

A more likely mechanism for populating the *D*-region with vibrationally excited O_2 is the following. McGrath and Norrish[7] in their studies on the photolysis of ozone observed vibrationally excited ground electronic state O_2 having up to 17 vibrational quanta as a result of the reactions

$$O_3(^1A) \longrightarrow O_2(^1\Delta_g) + O(^1D) - 91 \text{ kcal/mole} \qquad (2)$$

$$O(^1D) + O_2(^1A) \longrightarrow O_2^*(^3\Sigma_g^-) + O_2(^3\Sigma_g^-) + 138 \text{ kcal/mole} \qquad (3)$$

Absorption spectra[7] of the O_2^* indicated a maximum in intensity at about $v''=14$. Reaction (2), of course, refers to photolysis of O_3 in the region of the Huggins and Hartley absorption bands. We note that photo-dissociation resulting from absorption in the Schumann-Runge continuum would provide an additional source of $O(^1D)$ atoms, where

$$O_2(^3\Sigma_g^-) + h\nu \longrightarrow O(^3P) + O(^1D). \qquad (4)$$

Although no direct experimental observations have been reported for (1), it may be inferred from the O_2 density determinations of the *D*-region made by Jursa *et al.*[1] The O_2 density was calculated from the Schumann–Runge absorption bands in the region 1750–2000*A* obtained in their flight spectrograms. It is clear that this method of analysis would detect essentially only $O_2(^3\Sigma_g^-)_{v''=0}$ molecules. On the other hand, observations in the above spectral region would not detect $O_2(^3\Sigma_g^-)_{v''>0}$. In fact, the absorption spectra of the latter would only appear in the near ultra-violet region (2500–4000*A*) or in the region of Lyman-α (1216*A*). Now, the calculated O_2 density reported by Jursa *et al.* was consistently lower by a factor of 3·0–4·9 than those quoted by Miller (see Table 2 of Ref. 1). If it is assumed that Miller's values of the O_2 density in the *D*-region are correct, it is possible that an appreciable fraction of the O_2 were vibrationally excited, although nothing can be said about the distribution of vibrational excitation among these molecules.

It was further reported by Jursa *et al.* that the Lyman-α emission line did not appear at all in their spectrograms. They attribute this to the long absorbing path of O_2 despite the small absorption coefficient at this wavelength (0·27 cm^{-1}). This explanation appears to be contradictory since they were able to record photographically the (2, 0) to (13, 0) Schumann–Runge bands. Now the absorption coefficients of some of these bands are not much different from that at Lyman-α, implying, perhaps, that Lyman-α should have appeared on their spectrograms. It is, therefore, suggested that vibrationally excited O_2 could have completely absorbed Lyman-α in accordance with the process (1).

We suggest that an experimental test of the hypothesis presented above would be the measurement of the photoionization cross-section of vibrationally excited O_2 in the region of Lyman-α. It is possible that this may be considerably larger than that for NO.

EDWARD C. Y. INN

Lockheed Missiles and Space Division, Sunnyvale, Calif., U.S.A.

References

1. A. S. JURSA, Y. TANAKA and F. LEBLANK, *Planet. Space Sci.* **1**, 161 (1959).
2. M. NICOLET, *The Earth as a Planet* (Edited by G. P. Kuiper) p. 691 University of Chicago Press (1954).
3. V. BLACKMAN, *J.Fluid Mech.* **1**, 61 (1956).
4. K. WATANABE and F. F. MARMO, *J. Chem. Phys.* **25**, 965 (1956).
5. W. D. MCGRATH and R. G. W. NORRISH, *Nature, Lond.* **180**, 1272 (1957).
6. C. E. TREANOR and W. H. WURSTER, *J. Chem. Phys.* **32**,758 (1960).
7. W. D. MCGRATH and R. G. W. NORRISH, *Proc. Roy. Soc. A* **254**, 317 (1960).

SELECTED PAPERS TRANSLATED FROM

ARTIFICIAL EARTH SATELLITES U.S.S.R.

ISKUSSTVENNYE SPUTNIKI ZEMLI

(PUBLISHED BY THE U.S.S.R. ACADEMY OF SCIENCES, MOSCOW)

Translations of any papers abstracted here may be obtained from Pergamon Institute at a reasonable cost. Enquiries should be addressed to the Administrative Secretary, Pergamon Institute, at either Headington Hill Hall, Oxford, England, or 122 East 55th Street, New York, N.Y., U.S.A.

Papers marked * have been selected for publication in future issues of *Planetary and Space Science*.

Planet. Space Sci. Pergamon Press. 1961. Vol. 5, pp. 79-85. Printed in Great Britain.

ABSTRACTS

No. 1 (1958)

S. N. VERNOV, N. L. GRIGOROV, IU. I. LOGACHEV and A. E. CHUDAKOV: **The measurement of cosmic radiation on an artificial earth satellite** (pp. 5–8).

Cosmic ray intensities were measured over the altitude range 225 to 700 km, latitudes 40 to 70°N and longitudes 20 to 140°E, using a counter 100 mm long and 18 mm diameter; the average amount of material surrounding the counters was 10 g/cm². Two counters, both mounted in Sputnik II, were used, giving identical results throughout. The intensity at 225 km reaches a maximum of 36 counts/second at 65°N, and is about half this at 44°N. Intensity increases gradually with altitude, and at 700 km is about 40 per cent greater than at 225 km, for all latitudes. There is no marked or regular variation with longitude. Sporadic "flares" were observed, giving up to twice the normal counting rate over short periods.

M. L. LIDOV: **The determination of the density of the atmosphere from the drag observed on the first artificial earth satellites** (pp. 9–20).

Attempts to determine H, the height of the homogeneous atmosphere, from the drag experienced on artificial earth satellites near the perigee, were vitiated by the major diurnal variations in the state of the atmosphere. It was however possible to find $\rho\sqrt{H}$, where ρ is the density at perigee altitude; at 228 km the value is $1{\cdot}7\times10^{-6}$ kg/m², km$^{\frac{1}{2}}$ $\pm$ 40 per cent, i.e. 5–10 times greater than the value quoted on the basis of earlier rocket experiments. Data on the American satellite 1958α give a value of $0{\cdot}10825\times10^{-6}$ kg/m² km$^{\frac{1}{2}}$ at 368 km altitude. Availability of these values for 228 and 368 km makes it possible to tabulate the ranges of ρ and H consistent with both sets of data. Thus, ρ must lie between 3·2 and $2{\cdot}38\times10^{-13}$ g/cm³ at 228 km and between 1·37 and $1{\cdot}51\times10^{-14}$ g/cm³ at 368 km. The diurnal variations are considered and tentative values given for the diurnal range of $\rho\sqrt{H}$, namely from 2×10^{-6} kg/m² km$^{\frac{1}{2}}$ at noon to $1{\cdot}2\times10^{-6}$ kg/m² km$^{\frac{1}{2}}$ at midnight for altitude 228 km.

P. E. EL'IASBERG: **A determination of the density of the upper atmosphere from the secular variations of the orbital elements of the first two artificial earth satellites** (pp. 21–24).

Atmospheric density determinations were carried out for the perigee altitude (circa 225 km) of Sputniks I and II and the carrier rocket of Sputnik I, by a method involving representing the magnitudes of the secular variations of the orbital elements as a series of Bessel functions of an imaginary argument. Assuming Mitra's model of the upper atmosphere, the values obtained range from 2·8 to $4{\cdot}2\times10^{-13}$ g/cm³ at 225 km.

V. V. BELETSKII: **Motion of an artificial earth satellite relative to its centre of mass** (pp. 25–43).

The motion of an artificial earth satellite about its centre of mass, under the action of aerodynamic and gravitational perturbations, is investigated mathematically, taking into account the influence of retrogradation of the orbit due to the oblateness of the earth. Generalized equations are given, but for the purposes of analysis it is assumed that the satellite is dynamically symmetrical. The motion of a satellite is made up of regular precession (if dynamically symmetrical) about the vector of the kinetic moment, plus a secular precession and nutation of the same vector due to the perturbations. The secular motion takes place about an axis varying with time due to the retrogradation. If the influence of aerodynamic perturbations is the major one, the axis is close to the tangent at the perigee. If the magnitude of the gravitational perturbations exceeds that of the aerodynamic, as for satellites with perigee altitudes upwards of 500 km, then the trajectories of the kinetic moment vector may be of two main types, depending on the initial data. In the first, they are symmetrical relative to the plane of the orbit, concentric with the aerodynamic precession and flattened due to the gravitational perturbation. In the second, their centre lies close to the centre of precession. Motion relative to a retrograding orbit differs little from that relative to a fixed orbit.

*I. S. SHKLOVSKII: **Optical methods of observation of artificial earth satellites** (pp. 44–49).

Accurate determinations of the tracks of artificial earth satellites against the background of fixed stars are important for determining the orbit, which in turn yields valuable information on atmospheric resistance, including tidal currents, and on the shape of the Earth. This paper discusses the practical problems involved in optical observations, associated with the brightness of the satellite in comparison with that of the stars and the twilight glow conditions under which visual observations must be made, and with the accurate determination of the time of the observation. One method of timing is to scratch a "slit" on the photographic plate, so that the satellite itself illuminates a photocell to give a time reading. To increase the apparent brightness, it is suggested to make the satellite polyhedral, so that the plane faces give brilliant flashes as the satellite rotates. Schmidt-type optical systems are favoured for their high resolving power.

V. A. KOTEL'NIKOV, V. M. DUBROVIN, V. A. MOROZOV, O. N. RZHIGA and A. M. SHAKHOVSKII: **Use of the Doppler effect to determine the orbital parameters of artificial earth satellites** (pp. 50–61).

Mathematical expressions are derived for the variation

in received frequency due to the Doppler effect, under conditions of uniform rectilinear motion, uniformly accelerated rectilinear motion and motion on a curved trajectory. The methods of determining the instant of nearest approach, the distance of the satellite at this point and its velocity are discussed. Sputniks I and II carried a 40 megacycle transmitter, the signals from which were mixed at the receiving station with a standard frequency from a quartz oscillator, to produce beats for recording on magnetic tape. Good readings were obtained at distances up to 1000 km. The time of nearest approach was determined to within $\pm$ 1 sec at this range, the orbital velocity to within $\pm$3 per cent and the distance at nearest approach to about the same accuracy. From these first experiments, improvements have been possible in the equipment to improve the timing of the frequency variation curves and to include automatic frequency metering gear. These will reduce the errors to some hundredths of a second.

K. I. GRINGAUZ: **Rocket measurements of electron concentration in the ionosphere, using an ultra-shortwave dispersion interferometer** (pp. 62–66).

After a brief discussion of the shortcomings of reflected-signal radio-sounding of the ionosphere, a description is given of Soviet rocket experiments, culminating in the ascent of 21 February 1958, which attained an altitude of 473 km. The rockets carried two VHF transmitters operating on two different frequencies, one a multiple of the other. The phase and amplitude differences of the signals received at two different points were continuously recorded; simultaneously, the rockets were tracked optically and by radar and ionospheric radio-sounding records were made at nearby stations for comparison with the transmitted signals. It is concluded, in agreement with published American findings, that there is no sharply defined *E* layer below the maximum of the *F* layer. In contrast to the American work, however, it is found that ionization does not fall rapidly above the *F* layer maximum, but is still pronounced up to the highest altitude reached. Comparison of the rocket results with those of radio-sounding is useful in interpreting the latter. Thus it was found that the height within the *F* layer from which the radar waves are actually reflected is some 50–150 km lower than the apparent altitude recorded at the station, depending on the state of the ionosphere.

L. A. ZHEKULIN: **The altitude distribution of electron concentration from experiments with rockets and artificial earth satellites, and its influence on the propagation of radio waves** (pp. 67–79).

A mathematical study is made of the propagation of electromagnetic waves through the ionosphere, using the latest experimental data on ionization distribution. The solutions are expressed in the form of Airey functions. According to earlier hypotheses based on the existence of sharply defined ionized layers, reflection jumped from the *E* layer to the *F* layer on reaching a certain critical frequency. It is now known that the *E* layer corresponds to a sharp change in ionization gradient, and the critical factor is the frequency at which the dielectric permeability, corresponding to the ionization density at the level where the gradient changes suddenly, reverts to zero. On passing through this frequency, the signal delay suddenly increases so that the apparent reflection altitude thus increases suddenly. It is shown that when absorption in the ionized medium is small, the results obtained by the present solution approximate to those derived from the principles of geometrical optics, ionization gradient being equivalent to refractive index.

V. N. CHERNOV and V. I. IAKOVLEV: **Scientific experiments on the flight of an animal in an artificial earth satellite** (pp. 80–94).

A detailed account is given of the series of experiments culminating in the flight of the dog Laika in Sputnik II. The dogs must be systematically trained to endure the various conditions of long-period confinement, first under normal terrestrial conditions prior to take-off, secondly under the noise and acceleration of launching and finally under prolonged weightlessness in orbit. All these conditions, except the last, were reproduced singly and in combination in the laboratory and the training continued until the animals quickly reverted to the normal physiological state on cessation of the special conditions. The sealed container, its heating and ventilation system and the various pieces of experimental gear are described and illustrated, but not in great detail. Laika was fed on a mixture of rusk, dried meat and beef fat, ground up and mixed to a jelly-like consistency with water and agar-agar and was free to move in its harness over a short distance, so that measurements could be made of its locomotive activity under the various sets of stimuli. The signals transmitted to earth covered breathing and heart action, body temperature, movement and maximum pressure in the carotid artery, removed for the purpose into a special skin graft and constricted by a manometric sleeve. No substantial changes were recorded in the physiological state of Laika during seven days orbital flight. The satellite passed through a meteorite zone, but the cabin remained sealed and functioned normally throughout. There was no evidence of any effects due to cosmic radiation, but a final judgement on this would necessitate prolonged post-flight investigations.

No. 2 (1958)

*L. I. SEDOV: **Dynamic effects in the motion of artificial earth satellites** (pp. 3–9).

After a review of the various influences causing perturbations in the orbital and rotary motion of artificial earth satellites, due to atmospheric resistance the oblateness of the earth, the earth's magnetic field,, meteorite collisions and solar and lunar gravitation, the results are given of the full computation of observational data on the first three Soviet satellites. Graphs are given of the secular variations in period and in altitude at perigee and apogee, and the main data are summarised in

the following table, reproduced in full, referring to the immediate post-launching period in each case:

Satellite	Sputnik I	Sputnik II	Sputnik III
Period of rotation, minutes	96·17	103·75	105·95
Altitude at perigee, km	226–228	225	226
Altitude at apogee, km	947	1671	1881
Inclination of orbit, degrees	65·129	65·310	65·188
Daily variation in longitude of ascending node, degrees	3·157	2·663	2·538
Daily variation in perigee angle, degrees perigee angle,	0·432	0·407	0·326
Daily variation in period, sec/day	1·8	3·08	0·75

From these data, using accepted models of the atmosphere, it is calculated that the density at 225 km is between 3 and 4×10^{-13} g/cm^3.

E. G. Shvidkovskii: **Some results of measurements of the thermodynamic parameters of the stratosphere using meteorological rockets** (pp. 10–16).

Preliminary results are given of Soviet IGY experiments involving rocket probes of the stratosphere (up to 50 km) from the island of Kheis (Franz-Josef Land) and from mid-latitudes of the Soviet Union. The Antarctic measurements carried out at the same time, i.e. summer 1957 to spring 1958, are not yet available. Calculations have shown that the combination of temperature, pressure and altitude corresponds to thermodynamic equilibrium, so that in the majority of probes it is unnecessary to track the rocket, reliance being placed on temperature and pressure readings to find the corresponding altitude. Monthly temperature distribution curves for the two locations show that the isothermal layer is deeper (up to 26–32 km) in the Arctic than at mid-latitudes (21–26 km), and that in the positive gradient region above this, the higher temperatures are reached in the Arctic. Graphs of the temperature at various heights against time of year show maxima in the summer months; at 50 km the maximum coincides with the summer solstice, but with decreasing altitude it is later, reaching mid-July at around 30 km. The seasonal variations are more marked in the Arctic, where the stratosphere as a whole tends to be slightly warmer than at mid-latitudes.

B. A. Mitrov: **Perturbations of the gas medium caused by the passage of a satellite** (pp. 17–25).

Measurements of the density, ion content and concentration of the atmosphere are affected by the fact that the satellite is moving through the medium at a high velocity, compared to the average thermal velocity of the molecules, and also that adsorbed or leaking molecules are continually leaving the satellite surface and interacting with the medium. In addition, impacts occur between molecules of the medium and the satellite surface. The effects of these various molecular currents on the parameters of the medium in the vicinity of the satellite are considered mathematically, and it is shown that provided the instruments are carried on slender rods at an adequate distance from the main body of the satellite, the results should be satisfactory. This is confirmed by actual results (see papers by Mikhnevich and Istomin in the same issue). The presence of water vapour has been traced to volatilization from an adsorbed layer on the satellite, but no explanation is available for the fact that some of the water molecules were ionized. This could not be due to mere collisions, and points to the existence of a new and powerful radiation in the upper atmosphere.

V. V. Mikhnevich: **Preliminary results on the determination of atmospheric density above 100 km** (pp. 26–31).

After a very brief description of a magnetic manometer for pressures in the range 10^{-5}–10^{-7} mm Hg, and an ionization manometer for the range 10^{-5}–10^{-9} mm Hg, a few preliminary data are given of high-altitude density measurements made with their aid. These results agree in general with those quoted from American sources. In connexion with the satellite measurements, an interesting graph is given showing the variation in apparent pressure recorded as the satellite rotated. It is also pointed out that the gauge was "open" to the escaping adsorbed molecules from various parts of the satellite, and the minimum pressure readings recorded during the first three days of flight fell from 10^{-7} to 10^{-9} mm Hg as the satellite surface was gradually out-gassed in the rarefied upper atmosphere.

V. G. Istomin: **Investigation of the ionic composition of the earth's atmosphere by rockets and satellites** (pp. 32–35).

After a brief description of the equipment, a modified Bennett-type mass spectrometer covering the range 6–50 a.u., and the technique used, preliminary data are given on rocket measurements between 105 and 206 km and on Sputnik III. In the rocket experiment, ions of mass number 16 (atomic oxygen) and 30 (nitrous oxide) were found in the range 105 to 190 km, with a well-defined peak around 143 km, together with ions of mass number 16 (atomic oxygen) between 192 and 206 km. According to the data transmitted from Sputnik III, atomic oxygen is the main ionized constituent between 230 and 885 km, followed by atomic nitrogen in amounts of 3 to 8 per cent of the atomic oxygen ions; the proportion of nitrogen ions increases with increasing altitude. During the first few orbits, ions of mass number 18 (H_2O^+) were recorded, but gradually disappeared. Mitrov (paper in same issue) attributes this to adsorbed water molecules escaping from the satellite surface.

*V. I. KRASOVSKII: **Soviet investigation of the ionosphere with the aid of rockets and artificial earth satellites** (pp. 36–49).

A review is given of the picture of the ionosphere which is gradually emerging from Soviet investigations on four main lines, namely radio-sounding by radar techniques from ground stations, study of reception from extra-terrestrial sources (the sun and radio-stars, lunar radar), reception from rockets and satellites and finally direct measurements from instruments carried on rockets and satellites. No experimental details are given, except of the positive ion trap carried on Sputnik III. The constitution of the ionosphere, from present information, is as follows. There is one main maximum ionization level, corresponding to the *F*2 layer, roughly about 300 km in altitude, i.e. from 50 to 150 km lower than its apparent altitude from radio-sounding experiments. Below this maximum, the ion concentration increases steadily, with subsidiary peaks and has no sharply defined layers. Above the maximum, the concentration falls off much more slowly, and extends to very considerable altitudes (e.g., a value is quoted of $1{\cdot}9\times10^5$ positive ions per cm^3 at 795 km). There is no thermodynamic equilibrium between the electrons, ions and other particles in the ionosphere, and the effective electron temperature is much higher than that of the particles.

*S. SH. DOLGINOV, L. N. ZHUZGOV and N. V. PUSHKOV: **A Preliminary report on geomagnetic measurements carried out on the third Soviet artificial earth satellite** (pp. 50–53).

For the geomagnetic measurements required as part of the IGY programme, Sputnik III carried a magnetometer with magnetically saturated detector units, which set themselves automatically in the direction of the field vector. Due to the rotation of the satellite, this resulted in periodic variation of the interference due to other equipment on the satellite; thus the magnetometer also provided information on the rotational motion. Very little of the data transmitted back from the satellite has yet been processed, but preliminary indications are that the intensity curves agree well with those for the corresponding great circles on the earth's surface. Major variations, of a brief and sudden nature, were experienced as the satellite passed through the ionosphere, but as yet no interpretation is offered for this observation.

O. D. KOMISSAROV, T. N. NAZAROVA, L. N. NEUGODOV, S. M. POLOSKOV and L. Z. RUSAKOV: **The investigation of micrometeorites by rockets and satellites** (pp. 54–58).

A detailed description is given of the micrometeorite impact gauges carried on Sputnik III, their design and calibration and the method of processing the data for radio transmission. The gauge consists of an impact plate about 850 cm^2 in area, carried on an annular flat spring which on flexing distorts a number of piezo-elements made from ammonium phosphate. The damped oscillations are passed into a sorting circuit, which is then automatically closed for a period of 0·6–0·8 sec, giving a resolution of 12–17 impacts per second. The impulse could be measured over a linear range from 0·1 to 1000 gcm/sec. According to preliminary data, the average micrometeorite density encountered by Sputnik III was $1{\cdot}7\times10^{-3}$ impacts/m^2 sec, corresponding to $4{\cdot}4\times10^{-12}$ g/m^2 sec, but at one stage the satellite passed through a dense stream, in which the counting speeds rose to 22 impacts/m^2 sec at an altitude of 1700–1880 km, 10 at 1300–1500 and 9 at 500–600. Assuming that the impulse recorded is proportional to the energy of the incident particle, the average energy is about 10^4 ergs.

*V. I. KRASOVSKII, IU. M. KUSHNIR, G. A. BORDOVSKII, G. F. ZAKHAROV and E. M. SVEMLITSKII: **The detection of corpuscles with the aid of the third Soviet artificial Earth satellite** (pp. 59–60).

A brief preliminary note on the intense stream of corpuscular radiation encountered by Sputnik III on May 15, 1958 and recorded on two ZnS(Ag) fluorescent screens, shielded with aluminium foil of two different thicknesses. Of the three possible sources, namely ions (protons), X-rays or electrons, the latter is considered most likely, since the associated energy flux would then be the least. In this case, their energy was probably around 10^4 eV. The intensity was so great that the needle went off the scale, indicating at least 4×10^3 erg/sec cm^2 steradian. The possible origins of this radiation are considered; the risk to living beings from the X-rays excited from the satellite by such radiation is pointed out.

S. N. VERNOV, P. V. VAKULOV, E. V. GORCHAKOV, IU. I. LOGACHEV and A. E. CHUDAKOV: **A study of the soft components of cosmic rays beyond the atmosphere** (pp. 61–69).

Sputnik III carried a high efficiency photon counter, consisting of a cylindrical NaI(Tl) crystal, 40 mm in diameter by 39 mm long, and a photo-multiplier with a 40 mm diameter photocathode. The values recorded and transmitted by means of a duration-modulated transmitter were the counting rate for units of 35 keV liberated in the crystal, the photomultiplier anode current and the intermediate dynode current. From the last two recordings, it is possible to determine the total ionization in the crystal and its quantum composition. After describing the equipment in some detail, the paper deals with the problems of data processing, involving identification of the sources of excitation and particularly with the fact that the crystal continues to fluoresce some time after the energy source has been cut off, i.e. it exhibits memory effects. A high-intensity electron stream (10^3–10^4 particles/cm^2 sec steradian) occurs in the N polar region, but is not symmetrical with respect to the magnetic pole. The high intensities recorded in the equatorial region by van Allen *et al.* at altitudes around 1000 km, are confirmed, but little data is yet available from the records, since these were obtained from out-stations in the Southern hemisphere (the satellite reached apogee near the S pole). The equipment was not sufficiently sensitive

to detect direct solar gamma-radiation in the presence of the high secondary photon densities and the intense electron streams.

L. V. KURNOSOVA, L. A. RAZORENOV and M. I. FRADKIN: **Heavy nuclei in primary cosmic radiation** (pp. 70–74).

A description is given of the Vavilov–Cherenkov counter mounted on Sputnik III for determining the numbers of heavy atomic nuclei in primary cosmic radiation. The instrument had two threshold values, the lower being approximately 300 times the average impulse from a mu-meson, corresponding to $Z=15$–20, and the second at $Z=30$–40. The average counting speed for nuclei above the lower threshold was $1{\cdot}22\pm0{\cdot}08$ per minute. The instrument was incapable of actually counting particles above the upper threshold, but merely recorded a change if an odd number was encountered in one counting period. Only one such change was recorded, and it is estimated from probability theory that not more than one even number (i.e. a count of 2) could have occurred. Hence, over a period of 9 days, the number of such particles cannot have exceeded 1–3. This gives a ratio between the two groups of particle, $Z_{15-30}/Z_{>30}$ less than 0·01–0·03 per cent against the ratio of 30 per cent given by Yagoda (*Phys. Rev.*, **85**, 720, (1952)).

V. S. VAVILOV, A. P. LANDSMAN and V. K. SUBASHEV: **Solar batteries** (pp. 75–80).

The advantages of the use of silicon, with *p–n* junctions produced by phosphorus, boron or antimony diffusion, are discussed, and methods of improving their efficiency considered. The main batteries in Sputnik III were of this type and functioned perfectly for more than 6 months, supplying the multi-channel transmitting unit. When the satellite was in night, electro-chemical batteries were used, but it is eventually intended to use buffer-accumulators. Sputnik III also carried three small experimental solar batteries, one completely open, the second screened with polished glass and the third with frosted glass, in order to assess the effects of micro-meteorite impacts on the working surfaces. In fact, the erosion effect is so small that longer flights will be required to give quantitative results. Semiconductor devices are highly sensitive to temperature; special treatments (not described) have been devised which maintained steady temperatures of 16 to 30°C in the batteries on Sputnik III, and consideration is now being given to the possibility of using optical light-gathering systems to improve the outputs for satellites which can be automatically oriented on the Sun.

M. A. ISAKOVICH and N. A. ROI: **The acoustic method of measuring the mechanical parameters of meteorites** (pp. 81–82).

A brief note on the difficulties in measuring the mass and velocity, or any independent combinations of these parameters, of meteorites. Apart from difficulties associated with inability to predict at what point on the collector surface the impact will occur (non-central impacts causing a turning moment), the impact always produces a surface explosion in the collector material, and the energy of the ejected material exceeds that of the incident particle, i.e. the coefficient of restitution exceeds unity. The gauge described in the paper by Komissarov *et al.* in this issue is based on the hypothesis, due to Staniukovich, that in explosive impacts the net impulse received by the collector plate is a single-valued function of the kinetic energy of the incident particle. It also incorporates a multiplicity of piezo-elements to compensate for the effect of non-central impacts.

No. 3 (1959)

V. A. EGOROV: **Capture in the limited three-point circle problem** (pp. 3–12).

The limited three-point circle problem in celestial mechanics is the problem of the motion of a point mass of negligible mass under the influence of two large but finite point masses which are rotating about a common centre of mass at constant angular velocity. In the general three-point problem, the concept of capture means that if three point masses approach each other from infinite distances the distance between at least two of them will thereafter remain finite. Capture in the limited three-point circle problem means that the point of negligible mass approaching the two orbiting masses from infinity will thereafter remain at distances less than some finite limit. The problem is thus of interest in connection with artificial satellites in free flight after launching, in order to show whether it is possible to find trajectories according to which they can be put into permanent orbit about the moon or other planets, or whether they must necessarily return to orbit the earth or the sun. It is shown that if the ratio of the two attracting masses is sufficiently small, a point coming from infinity into the sphere of attraction of the smaller mass must leave that sphere again. In particular it is impossible for the moon to capture a missile in free flight launched from the earth; again, it is impossible for one of the planets to capture a missile from the sphere of attraction of the sun.

V. V. BELETSKII: **The libration of a satellite** (pp. 13–31).

An earlier paper by Beletskii ("Motion of an artificial earth satellite relative to its centre of mass,") *Iskusst. Sputniki Zemli*, 1958, No. 1, 25–43) discussed the motion of a satellite under the action of aerodynamic and other perturbations, assuming that the initial kinetic energy of the satellite about its centre of mass was large in comparison with the perturbing forces. The present paper discusses the case when the kinetic energy is small, in which case the satellite may librate about an equilibrium position. It is first shown that such an equilibrium position is possible for a body in circular orbit in a Newtonian central field of forces if and only if the major axis of the ellipsoid of inertia of the body is directed along the radius vector of the orbit and the minor axis normal to the orbital plane. The conditions under which

libration about the stable equilibrium position can occur. are explored mathematically, and it is shown that the initial angular velocities about the centre of mass must be small, i.e. comparable with orbital angular velocity if libration is to occur rather than the type of motion discussed in the earlier paper. The effects of the various perturbing influences are not dealt with, but it is stated merely that "investigation indicates that if the stability conditions and certain other natural conditions are satisfied, libration in the presence of the various perturbations will not differ greatly from the unperturbed libration".

V. F. Proskurin and Iu. V. Batrakov: **First-order perturbations in the motion of artificial satellites due to the oblateness of the earth** (pp. 32–38).

The effects of the oblateness of the earth on the motion of artificial satellites of any orbital inclination are developed on the assumptions that the oblateness can be represented as giving the earth the shape of a regular ellipsoid of rotation and that the degree of oblateness is small, so that when the perturbation functions are expanded as power functions, terms of the second and higher orders can be neglected. Having derived the necessary equations, a numerical example is given for a satellite with the orbital characteristics: major semi-axis 7286·88 km, eccentricity 0·099493, inclination 65·4900°. The altitude at perigee can vary by 4·2 km, whilst the angular characteristics of the orbit will be periodically perturbed to the extent of several minutes of arc.

Iu. V. Batrakov and V. F. Proskurin: **Perturbations of the orbits of artificial satellites due to aerodynamic resistance** (pp. 39–46).

A mathematical study is made of the first-order perturbations in the characteristics of the elliptical orbits of satellites due to atmospheric resistance. It is assumed that the earths atmosphere has a perfectly spherical density distribution and that the earth's attraction can be represented as that of a point of equal mass situated at the centre of inertia. It is shown that in addition to the secular perturbation, i.e. the retardation and spiralling effect on the orbit, there are short-period perturbations. Numerical calculations however show that whereas the shortening of the semi-major axis due to air resistance is 2·303 km/day at the initial value of 7286·88 km, the amplitude of the largest periodic term in the major semi-axis length is only about 50 m. Similarly, the periodic perturbations in the longitude at perigee are about half a second of arc. The periodic perturbations can therefore be neglected in processing the results of observations.

V. M. Vakhnin and V. V. Beletskii: **Observations on artificial satellites by the expectancy method** (pp. 47–53).

The paper describes a method of setting the instruments at observatories to give the maximum likelihood of picking a satellite up as soon as it appears above the horizon, so that accurate observations can be made and the orbital characteristics determined as quickly as possible. In the simplest case, the orbital plane can be regarded as stationary relative to the stars. Then, if the period of the satellite is an exact fraction of one sidereal day, after one day the satellite will reappear from exactly the same direction at exactly the same sidereal time. Conversely, if it does not describe a whole number of revolutions in the day, it will appear at a later time and from a more westerly point. Clearly, the delay and the westerly displacement are proportional; hence a curve can be constructed plotting direction of observation against time, and by traversing the instrument according to this curve it can be ensured that it is always pointing in the correct direction to pick the satellite up, whenever it should appear. The actual calculations are complicated by orbital perturbations, parallax and the precession of the orbit, but provided the visual angle of the instrument is adequate, only the latter effect need be taken into account.

P. E. El'iasberg: **Dependence of the secular variations of orbital elements on aerodynamic resistance** (pp. 54–60).

This paper is a further development of the idea first put forward by Iatsunskii (*Uspekhi fiz. nauk*, **63**, No. 1a 1957) of expanding the secular variations of orbital elements in a series the terms of which are Bessel functions of an imaginary argument. The formulae thus obtained are simple and readily understandable and can be used to solve a number of problems, such as the relationship between the secular variations in various elements and the magnitudes of those elements, the estimation of the density of the atmosphere from measurements of secular variation, etc., which are briefly touched upon and discussed in the light of the formulae obtained.

M. A. Lavrent'ev: **The problem of puncture at cosmic speeds** (pp. 61–65).

A brief mathematical note on the problem of the phenomena accompanying impact and puncture at relative velocities of 50–100 km/sec, i.e. about 10 times the relative velocities studied in ballistics research.

I. S. Shklovskii and V. G. Kurma: **Determination of the density of the atmosphere at 430 km by the diffusion of sodium vapour** (pp. 66–76).

The vapour cloud diffusion method, first proposed by Bates, offers a valuable independent check on density determinations derived from measurements of the orbital perturbations of artificial satellites. In the U.S.A. the method adopted was to release the sodium as the rocket ascended, giving a narrow strip of vapour; in the Soviet experiment, carried out on 19 September 1958, the rocket was at its apex and almost stationary when the sodium was released, at 430 km. A 2-kg charge of sodium, vaporized by thermite, was discharged horizontally from the stabilized rocket in a period of 10–20 seconds. The twilight zone was 300 km deep at the time of release, giving good visibility from a very wide area. The cloud was irregular at first, but soon took on a spherical shape,

and its total luminosity was constant from about 100 to 1000 seconds after release. The maximum diameter reached about 100 km, with the square of the effective radius proportional to time, thus showing the normal diffusion relationship. The density value calculated was $(6{\cdot}7 \pm 2) \times 10^{-15}$ g/cm^3, which agrees well with the recent series of determinations between 300 and 750 km.

I. M. Imianitov and Ia. M. Shvarts: **Methods of preventing interference currents at the input to electrostatic fluxmeters operating in conducting media** (pp.77–83).

In the electrostatic fluxmeter used on an isolated body in a plasma, the measuring plate is a small area of the body surface itself, insulated from the rest except via a resistance. By periodically screening this area, by means of a rotating screen a small distance above it, from the space charge which always surrounds a charged surface in a plasma, the surface charge is caused to flow to and from the measuring plate, and its magnitude can be measured by measuring the current through the resistance. Other currents can however be set up which interfere with that due to surface charge. They have the common feature that they are 90° out of phase with the signal current, but the use of a synchronous detector which would eliminate them is a practical impossibility. The alternative examined is to make the measuring plate and screen in the form of open grids of given electrical and optical transparency, thus reducing the modulation of the charged particle flux giving rise to the interference currents and the absolute magnitude of the flux incident on the measuring plate. Apparatus of this type was successfully tested in the plasma of a mercury discharge tube.

V. V. Mihknevich, B. S. Danilin, A. I. Repney and V. A. Sokolov: **Some results of determinations of the structural parameters of the atmosphere by means of the third Soviet satellite** (pp 84–97).

After a brief description of the manometric equipment carried on Sputnik III and the methods of evaluating the data, tables are given showing the values every 5 km from 225 to 500 km of the mean molecular weight, the number of particles per unit volume, the density, temperature and pressure of the earth's atmosphere. Useful comparison tables and graphs are also given of the data from Soviet and American rockets and satellites, both by manometric measurement and by calculation from satellite retardation. Diurnal variations and variations with geographical position are briefly discussed.

V. G. Istomin: **Radio-frequency mass-spectrometer for studying the ionic composition of the upper atmosphere** (pp. 98–113).

The Soviet results on the study of the ionic composition of the upper atmosphere (to 885 km) were reported by the present author to the 5th Assembly of the IGY Special Committee in Moscow in 1958. The present paper gives a very detailed description of the mass-spectrometer used, including an annotated circuit diagram, views of constructional details and a discussion of the operating technique and operating characteristics.

S. A. Kuchai: **Errors in pressure measurements due to minor leaks in the skin of a satellite** (pp. 113–147).

The pressure measurements carried out on Sputnik III were made by an instrument with a working range between 10^{-6} and 10^{-9} mm Hg. The effects of gases adsorbed on the outer surfaces upset the readings for a time but rapidly wore off. Minor leaks from within the satellite could however represent a source of continuing error. The possible magnitudes of the errors arising from this source are calculated.

Iu. V. Zonov: **Interaction between a satellite and the earth's magnetic field** (pp. 118–124).

Zonov discusses first the currents induced by the forward motion of a satellite through the earths magnetic field, secondly the retardation of its rotation about its own axis due to eddy current effects and thirdly the perturbing forces of the magnetic field on a non-rotating satellite. The flight of the satellite through the field produces a potential difference of some tens of volts from one side to the other, which can affect certain measurements. If the satellite is rotating, the presence of this potential difference corresponds to a current flow round the skin, but the retarding effect is quite negligible (velocity reduced by a factor of e in 10^7 years). Eddy currents have a much greater effect (angular velocity reduced by a factor of e in 70 days for a Sputnik I type satellite, 11 days for the avant-garde type). In the third case, when the satellite is not already rotating, the magnetic field of the earth can produce a perturbing influence, causing it to commence to rotate.

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 87–91. Printed in Great Britain.

THE SYSTEMATIC BEHAVIOUR OF HYDROGEN EMISSION IN THE AURORA—I

M. H. REES, A. E. BELON and G. J. ROMICK
Geophysical Institute University of Alaska College, Alaska

(*Received* 10 *June* 1960)

Abstract—Analysis of short duration spectra taken in rapid succession reveals a systematic behaviour of the aurorally associated hydrogen emissions. Two principal effects are noted. Firstly, there is a daily motion along the geomagnetic meridian in which the spectral line progresses from the northern horizon southward, followed by a return to the north some time during the night. Secondly, the limit of the southern extent of the hydrogen light correlates well with the magnetic activity, i.e. the latitude at which the protons penetrate to auroral heights is a function of the magnetic disturbance.

1. INTRODUCTION

During the winter observing season 1958–1959 a newly constructed meridian type spectrograph[1] was operated at College, Alaska (geomagnetic latitude 64.65°N). The chief advantage of this instrument is its speed which allows the use of a routine fifteen minute exposure. It well registers the major auroral emission features, including the hydrogen line H-alpha. This relatively high time resolution makes it possible to study the changing course of the various spectral features of the aurora along a magnetic meridian. The Balmer lines of hydrogen are particularly important to the excitation of the aurora. Their presence in the auroral spectrum has often been reported as being erratic. In this paper a recurrent behaviour pattern of the hydrogen H-alpha emission line is reported and discussed.

2. THE OBSERVATIONAL DATA

The spectrograph was operated continuously from 21 October 1958 through 24 March 1959. Excepting a period of two weeks (26 January–10 February) during which the instrument was oriented east–west, the spectrograph monitored a narrow band in the magnetic meridian. Fig. 1 illustrates the meridional progression of H-alpha recorded during the night 8–9 December 1958. The time intervals selected for Fig. 1 best illustrate the progression of the H-alpha line from north to south during the first part of the night, and south to north toward morning, however, spectra were obtained continuously at fifteen minute intervals from 1600 to 0700 150° WMT*. North is at the top of each spectrum, south at the bottom; a small portion in the zenith is occluded by the camera placed above the spherical mirror. The most prominent auroral emission features recorded on this high speed low resolution instrument are indicated on the spectra, although here we shall concern ourselves only with the H-alpha line.

The spatial distribution of the hydrogen H-alpha line within the field of view of College for six consecutive nights is shown in Fig. 2. The solid lines represent the extent of the emission line visible on the film while the circles indicate the position of an intensity maximum if such a maximum clearly exists. It should be noted that the absence of a line for a particular time does not necessarily mean that H-alpha was not present. During very intense periods of overall auroral luminosity the H-alpha line is obscured by the much more intense neighbouring 1st positive bands of nitrogen. We have a few cases of this brief period of absence which cannot be explained by this contamination effect. It is not unreasonable to infer from the overall picture that, very likely, the peak of the H-alpha emission had moved south of our

* 150° West Meridian Time (Alaska Standard) will be used throughout this paper.

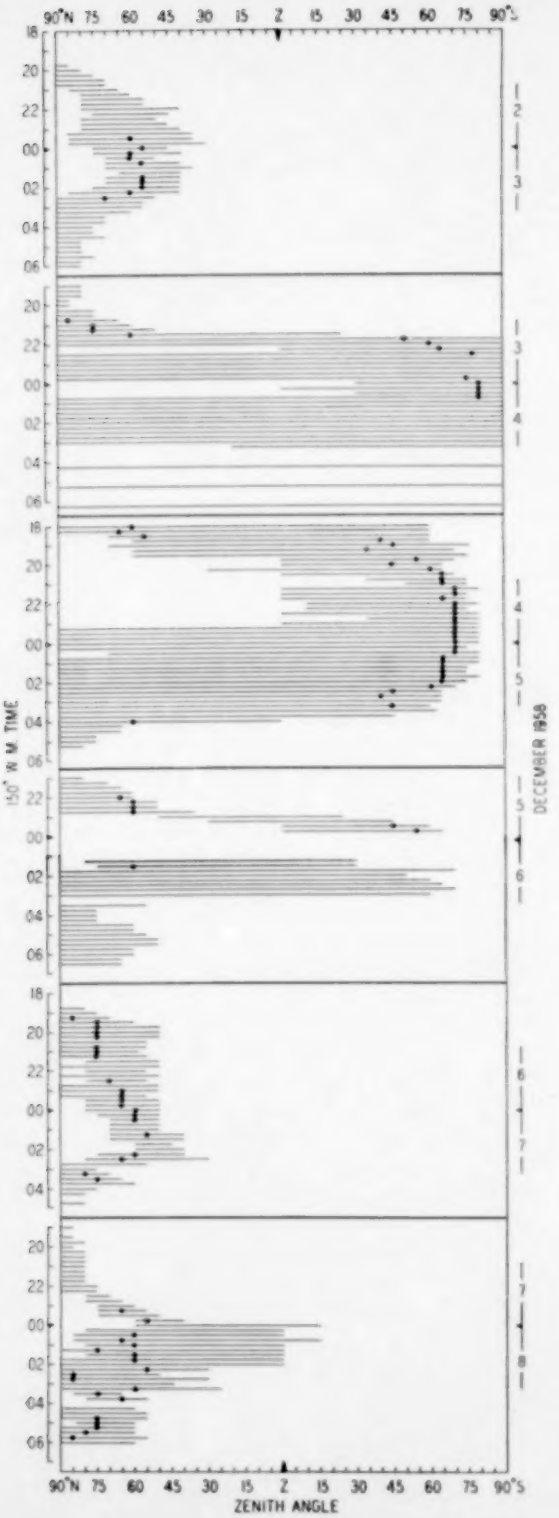


Fig. 2. Meridian extent of H-alpha line on the spectrum at 15 minute intervals throughout 6 consecutive nights. The circles indicate the position of the H-alpha intensity peak when this was clearly evident. See text for detailed explanation of the Figure.

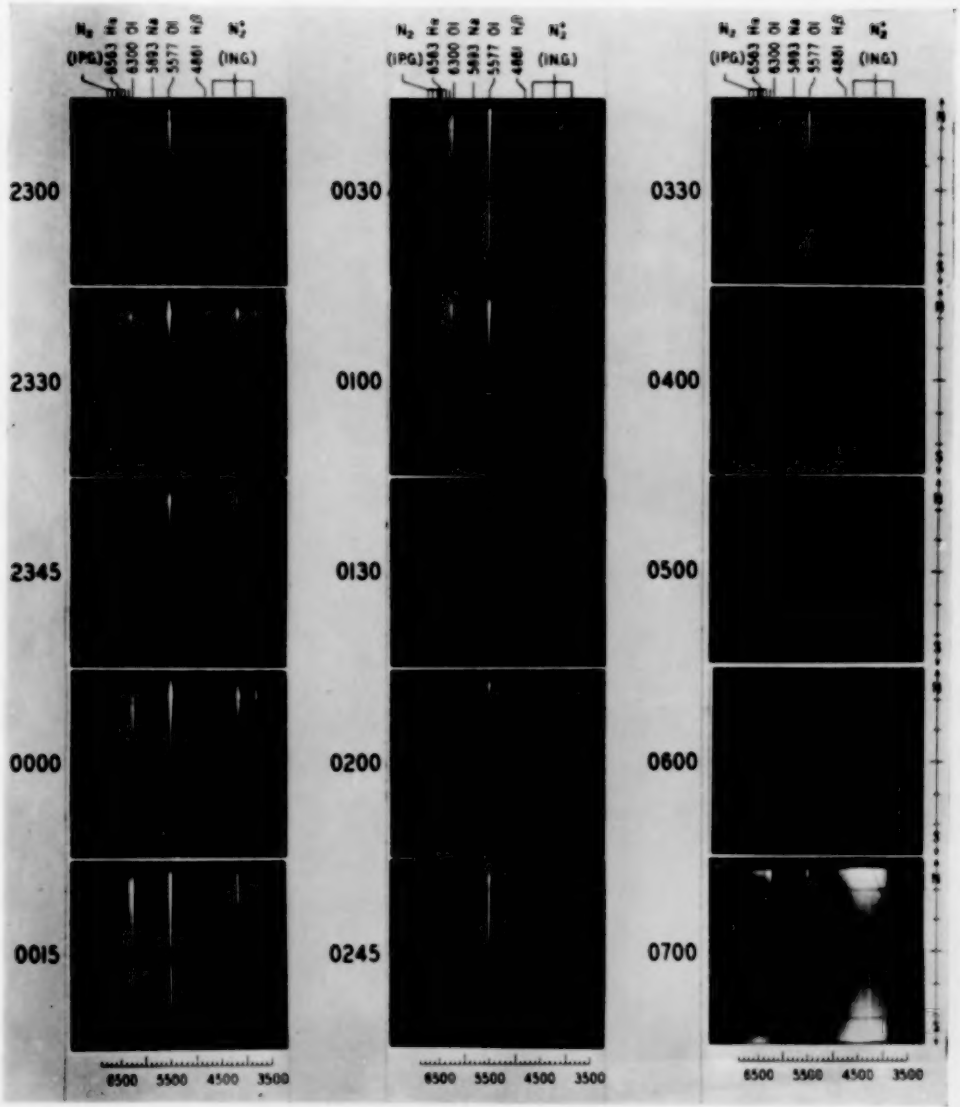


Fig. 1. (a) Selected meridian spectra for the night 8–9 December 1958. The time intervals (150° WMT) apply to the duration of the spectroscopic exposures.

(facing page 88)

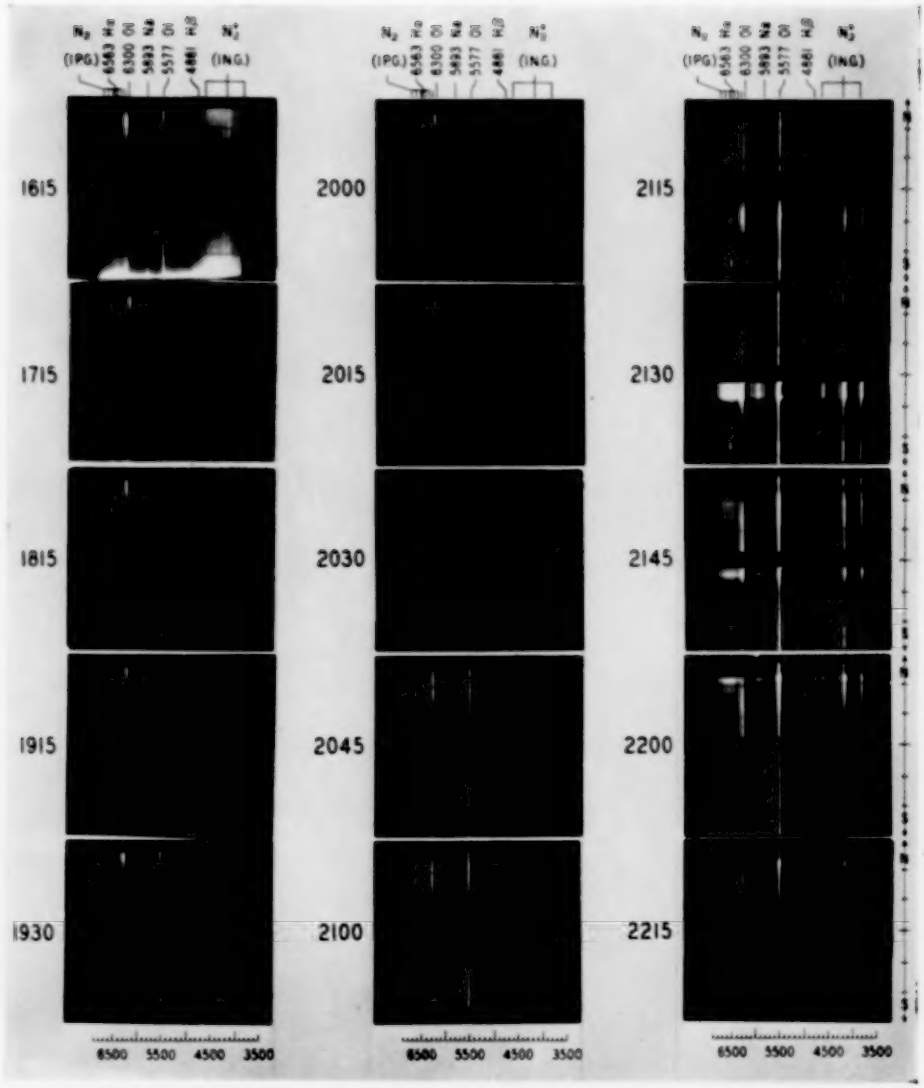


Fig. 1 (b). Selected meridian spectra for the night 8–9 December 1958. The time intervals (150° WMT) apply to the duration of the spectroscopic exposures.

station, out of the field of view of the spectrograph. The very rapid southward progression on 5–6 December (see Fig. 2) is indicative of this behaviour.

The observational evidence is summarized in the Table. The maximum southern extent has been divided, arbitrarily, into five categories corresponding to various zenith distances as derived from the spectra. Nights during which the H-alpha line, after extending to the southern horizon, was present on virtually all subsequent spectra over the entire meridian, and active nights characterized by H-alpha emission disappearing beyond the southern horizon and not returning while observations were in progress for the night, were put into the last category of the Table. It is noted that only two nights out of 62 usable ones showed no detectable H-alpha line. The following observations supplement the Table with regard to the behaviour of the hydrogen emission:

(a) During the two weeks when the spectrograph was aligned along geomagnetic east-west the appearance and disappearance of H-alpha was more or less simultaneous over the entire 180° field of view. This confirmed our conclusions from visual observations and all-sky camera pictures that the hydrogen emission is associated with a broad diffuse arc. It should be recalled, however, that a field aligned arc will not appear all at one time in the geomagnetic prime vertical at stations with a small dip angle.

(b) This arc associated with the hydrogen emission frequently is barely visible to the eye or the all-sky camera.

(c) H-alpha is definitely not associated with all auroral arcs. Inspection of over 2000 spectra showed many examples of arcs with and without the hydrogen emission, while the intensities of the green auroral line of oxygen were comparable. Also, hydrogen is not associated solely with quiet arcs. After an auroral display has reached full development hydrogen may occur in any type of auroral form, or in essentially formless aurora.

(d) The appearance of the hydrogen arc and its motion at least to the zenith *always* precedes the bright or breakup phase of an auroral display. In the majority of cases the H-alpha emission peak has moved well south of the magnetic zenith (at College) before breakup.

(e) Recession of the H-alpha arc may occur at any time after it has reached its southern extent. Frequently this coincides with the general recession of the aurora back to the north. However, when the aurora is active after breakup this recession may occur while the display is in full progress all over the sky. For

Table 1.
Maximum Southern Extent of H-alpha Emission in Terms of Zenith Distance of Observation Along the Magnetic Meridian.

Observational statistics	Number of nights
Total number of nights during which spectrograph produced meridian spectra, season 1958–59	133
Number of nights during which interference by moon and/or clouds made spectra unusable	71
Number of clear nights during which H-alpha was not detectable at all	2
Sample type of usable nights	
Maximum southern extent between 90°N and 45°N ...	9
Maximum southern extent between 45°N and zenith ...	8
Maximum southern extent between zenith and 45°S ...	13
Maximum southern extent between 45°S and 85°S ...	18
Remaining at maximum extent throughout observing time or extending beyond 85°S	12

this reason the recession is more difficult to trace than the progression.

3. DISCUSSION

Although protons may not be as important as electrons in the overall picture of auroral excitation it is nevertheless of interest to investigate the significance of the systematic progression of the hydrogen emission.

A general correlation between auroral and magnetic activity is well known in the literature[2]. However, few quantitative comparisons have been made due to the difficulty of establishing a measure, or criterion for "auroral activity". The relation of three-hour range index K to aurora seen at Ithaca, New York, has been studied by Gartlein[3]. Malville[4] has examined the daily variation in the latitude of the most northerly aurora observed visually, and with an all-sky camera at Ellsworth, Antarctica, as a function of the magnetic K index of Lerwick. Roach and Rees[5] have investigated the absolute intensity of the green auroral line with respect to the K index at College, Alaska. All the above cited studies show a positive correlation between the magnetic activity and the particular auroral criterion employed. Bartels and Chapman[6] find that the stronger the geomagnetic disturbance the farther the aurora advances towards the equator. Their conclusion is based on a comprehensive study of a 21-year series of geomagnetic activity data, represented in connection with the appearance of aurora in lower latitudes. In addition, Malville[4] found that the percentage of nights showing hydrogen emission on the IGY patrol spectra at Ellsworth increases with increasing Lerwick K index. From extensive spectroscopic measurements carried out in New Zealand, Sandford[7] concluded that the "forbidden" atomic oxygen emission lines [OI] tend to increase in intensity with magnetic disturbance.

The average magnetic activity (in γ) for each of the 60 nights referred to in the Table was determined by converting the 3-hour K-indices to the equipment range, ak, according to the co-ordination appropriate for College. We then computed the average magnetic activity for all nights falling into each category of southern extent of H-alpha. The result, shown in Fig. 3, indicates that the southern extent of the H-alpha emission line in the aurora can indeed be associated with the magnetic disturbance.

Magnetic activity in all latitudes, and especially in auroral regions has a clear daily variation[2]. We find that this is true also for

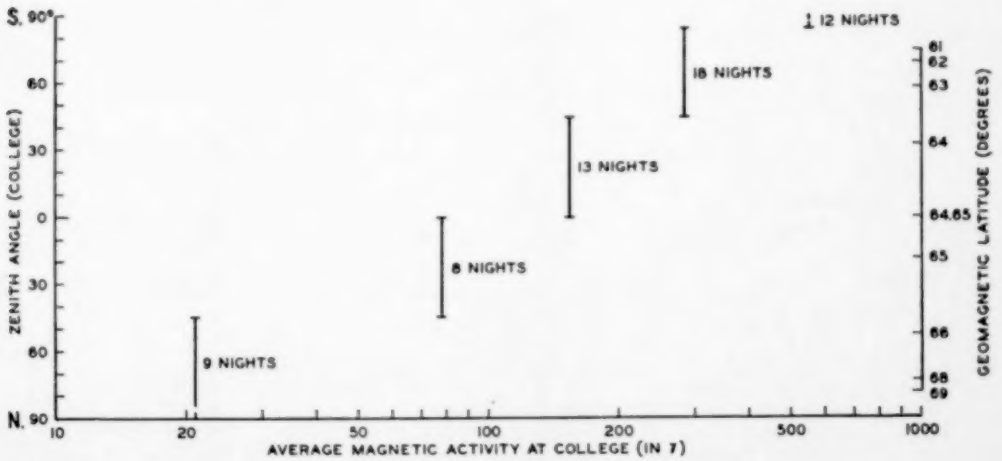


Fig. 3. Maximum southern extent of H-alpha emission as a function of magnetic activity. Averages for 60 nights. The geomagnetic latitude is computed under the assumption of a 100 km height.

the appearance of H-alpha, the southern extent of which is greater near the middle part of the night. The variation is similar whether the general level of magnetic activity is low or high. During magnetic storms the daily variation carries the H-alpha luminosity to numerically low latitudes. Fig. 2 shows an example of this, namely on the night of December 3–4. A magnetic storm began on 3 December 1958 at 1435 local time. For the 3-hour intervals 20–23, 23–02, 02–05 (150°WMT) on this night *K*–College had the values 5, 5, 7. Although very disturbed conditions prevailed for some time the diurnal variation of hydrogen is still clearly evident. Over the period of observations covered by this report there were five S.S.C. magnetic storms during which usable spectra were obtained. In all cases the extent of H-alpha fell into the extreme southern category, disappearing beyond the southern horizon or remaining at maximum extent throughout the night.

On the basis of the observations reported in the present paper, Rees and Reid[8] have advanced a semi-quantitative hypothesis attempting to explain the systematic behaviour of hydrogen in the aurora. Trapping of solar corpuscular radiation by the geomagnetic field is examined. The deformation of the geomagnetic field, a compression on the day side of the earth, is under solar control and remains more or less constant with respect to the sun while the earth rotates within this deformed pattern. The mean latitudinal motion of the hydrogen emission is attributed to a diurnal effect while the variations from the mean are thought to be due to the changing level of solar activity. These ideas are elaborated in Part II of this paper.

High time resolution studies of hydrogen should prove useful in extending our knowledge of the nature of the interaction between solar corpuscular streams and the earth's field. Unlike all the other auroral emission features hydrogen lines provide us with direct evidence that particles are penetrating into the earth's auroral zone ionosphere from outside, that the latitude of penetration is well correlated to the magnetic disturbance, and that the distortion of the field is under solar control.

Acknowledgement—The research reported in this paper has been supported by a grant from the National Science Foundation.

REFERENCES

1. K. C. Clark and G. J. Romick, *J. Opt. Soc. Amer.* **49,** 141 (1959).
2. S. Chapman and J. Bartels, *Geomagnetism* Oxford University Press (1951).
3. C. W. Gartlein, *Trans. Amer. Geophys. Un.* **25,** 533 (1944).
4. J. M. Malville, *J. Geophys. Res.* **64,** 1389 (1959).
5. F. E. Roach and M. H. Rees, *J. Geophys. Res.* **65,** 1489 (1960).
6. J. Bartels and S. Chapman, *Abhandlungen der Akademie der Wissenschaften in Goettingen Mathematisch-Physikalische Klasse*, No. 1 (1958).
7. B. P. Sandford, *Ann. de Géophys.* **15,** 445 (1959)
8. M. H. Rees and G. C. Reid, *Nature, Lond.* **184,** 539 (1959).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 92–98. Printed in Great Britain.

WAVE FORMATIONS IN NOCTILUCENT CLOUDS*

B. HAURWITZ

High Altitude Observatory, University of Colorado, Boulder, Colorado

(*Received* 23 *May* 1960)

Abstract—The implications of the appearance of wave and billow structure in noctilucent clouds are discussed. While in some cases the billows can be explained as effects of the angle under which the clouds are viewed, the wave motion itself with its varying horizontal convergence and divergence must also give rise to billows. In any case the presence of billow structure presents no difficulty to the hypothesis that the noctilucent clouds consist of dust particles. Observations of the wind and temperature distribution at the heights of the noctilucent clouds and during the time when waves are visible could give information about the dynamics of the atmospheric layers, in conjunction with observations of the lengths and velocities of these waves.

1. INTRODUCTION

Like the clouds in the lower atmosphere the noctilucent clouds near the mesopause show frequently a wave structure. Photographs of this phenomenon have been published by Witt[1] and others. For the clouds of the lower atmosphere the appearance of a wave structure, often referred to as billow clouds or mackerel sky, was first explained by Helmholtz as due to wave motions at an interface where wind and density change suddenly in the vertical direction. The vertical motion connected with the internal waves leads to condensation of water vapour where it is directed upward and to evaporation where it is directed downward; hence the array of the clouds in long parallel bands of equal distance. Where sufficient observational data concerning the wave length, the wind and temperature distribution of the atmosphere are available, the observed spacing and orientation of the cloud bands are in good agreement with the theoretical predictions, confirming Helmholtz's explanation[2, 3].

The noctilucent clouds, unlike the clouds of the lower atmosphere, are, according to many students of this phenomenon, made up of dust particles, presumably of cosmic origin (see for instance[4, 5]). But this explanation has so far not found general acceptance; many investigators regard the noctilucent clouds as formed of water substance, for instance[6, 7]. If the noctilucent clouds do consist of dust particles, their band structures can, of course, not be explained by condensation resulting from wave motions at an interface. In fact, it has occasionally been suggested that the appearance of the billow structure in noctilucent cloud displays poses a serious obstacle to the hypothesis that they consist of dust particles.

This problem has been discussed briefly by Ludlam[5]. He suggested that the appearance of billows in noctilucent clouds may be due to wave motions of the cloud layer. For an observer looking along low angular elevations, adjacent lines of sight will pass through layers of appreciably different optical thickness so that the cloud sheet will show a band structure. The two requirements for this explanation to hold are (1) that the noctilucent clouds are low on the horizon and (2) that the wave crests and troughs are normal or nearly normal to the line of sight. The first requirement is apparently generally satisfied. A few earlier observations of noctilucent clouds extending high into the sky and even through the zenith, quoted by Vestine[8], must be regarded as doubtful in view of the difficulty to account for their illumination by the sun at such high elevation angles. No systematic studies are available with respect to

* The research reported in this document has been sponsored by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command under contract AF 19 (604) 5488.

the second requirement, viz. the orientation of the billows with regard to the line of sight. But photographs published by Witt[1] show billows in directions quite far from the normal to the line of sight. In such cases it does not seem possible to explain the billow structure in terms of the viewing angle.

Before an alternate or possibly complementary explanation for the appearance of the billow structure in terms of the hydrodynamics of the wave motion is given, a rough quantitative estimate of the effect of the differences in viewing angle will be made for comparison with the explanation to be offered here.

To obtain such an estimate the clouds may be approximated by segments of straight lines, as shown in Fig. 1. While this approximation

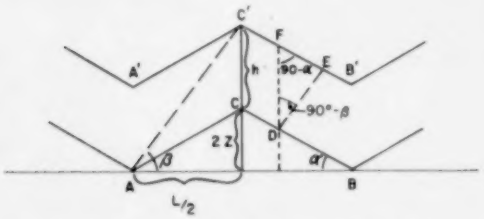


Fig. 1. Estimate of the effect of the viewing angle on the appearance of the cloud.

represents a considerable idealization of the appearance of the clouds, it will give at least a quantitative estimate. It must also be pointed out that exactly sinusoidal waves may by no means be the best approximations to the actual form of the cloud waves.

The distance L between A and B, the two lowest points of the lower cloud base, corresponds to the wave length. The distance $2Z$ from AB to C, the highest point of the lower edge, approximates twice the wave amplitude. The distance CC′, or AA′ and BB′ may be referred to as the thickness of the cloud.

Suppose that the observer looks along AC′, to select a favourable direction for sighting through much cloud material. Then

$$\tan\beta = 2\,(h+2Z)/L$$

Further,

$$AC' = (h+2Z)/\sin\beta$$

When looking through the descending part of the cloud, the line of sight will be in a slightly different direction from AC′. But as long as the distance of the cloud from the observer is large compared to the wave length, the difference may be neglected, and DE through this part of the cloud may be assumed parallel to AC′. If α represents the angle between the cloud base and the horizontal

$$\tan\alpha = 2\frac{2Z}{L}$$

This quantity might be called the steepness of the wave (although in oceanography a quantity corresponding more closely to half this value is defined as the steepness). If FD is normal to AB, the angles of the triangle DEF are as indicated in Fig. 1 and

$$DE/h = \sin(90-\alpha)/\sin(\alpha+\beta)$$

Hence

$$DE/AC' = h\frac{\cos\alpha}{\sin(\alpha+\beta)}\frac{\sin\beta}{h+2Z} =$$

$$= \frac{h}{h+2Z}\frac{1}{(\tan\alpha/\tan\beta)+1}$$

or

$$DE/AC' = h/(h+4Z) \qquad (1)$$

for the desired ratio. In general the amplitude Z will not be large compared to the cloud thickness, h, and the ratio will, therefore, also not be very small. For instance, if $h=4$ km, $Z=1$ km, which are the values chosen in Section 4, it would be only one half.

Other estimates of the effect of the viewing angle based on different directions, for instance, that parallel to the cloud base AC, do not give significantly different results.

The foregoing is admittedly a rather crude estimate of the effect of sighting on the appearance of the clouds, but it shows that the differences in brightness may not be as large as one might expect at first. A comparison with the effects due to the wave motion which are discussed in Section 4 will show that the two effects are comparable in magnitude and may supplement each other.

The discussion of the wave motion and its effect on the noctilucent clouds will also show

which atmospheric parameters are important for the dynamics of this motion. Accordingly some suggestions are offered for observations which might lead to a better understanding of the dynamics of these layers of the atmosphere.

2. WAVE MOTION AT AN INTERFACE

In order to simplify the mathematical discussion, it will be assumed that the atmosphere is incompressible and homogenous, except for the density discontinuity at the interface. The assumption of two incompressible and homogeneous layers neglects the inherently stable stratification of the atmosphere and leads, in the case of the billow clouds, to wave lengths which are shorter than those observed[2]. But since no adequate data are at present available for a comparison of the observed wave lengths in noctilucent clouds with the theory, a more accurate formula for the wave length derived for a compressible fluid need not concern us here. Because of the height of the noctilucent clouds above the ground, the two fluid layers can be regarded as infinitely deep even though the lower layer is bounded by the surface of the earth.

It will also be assumed for the sake of mathematical convenience that sharp discontinuities of wind and density or, what is the same because of the continuity of the pressure at the interface, of the temperature exist at the interface. In the atmosphere such sharp discontinuities are not found, but the difference between a fluid model with sharp discontinuities and the real wind and temperature distribution with a narrow zone of transition is minor. Sekera[9] has studied the effect of a zone of wind shear on the wave length of billow clouds in detail and has shown that its effect is to increase the wave length, in the maximum by a factor 2, but by a considerably smaller amount when the thickness of the shear zone is decreased. For instance, if the wave length in the case of a sharp discontinuity would be 3000 m, a shear zone of 1000 m thickness would increase it by only 25 per cent. Under these circumstances the idealization of the transition zones of wind and temperature changes to sharp discontinuities can be retained for the purposes of the present paper.

We assume first that the winds in both layers are in the same or opposite directions and choose a co-ordinate system, fixed relative to the earth's surface, whose z-axis points vertically upward. This case has been discussed by Lamb[10]. The wave motion may be regarded as two-dimensional, that is independent of the y co-ordinate. This simplification implies an infinite lateral extent of the wave, an assumption which does not give rise to a serious error since the wave crests are as a rule much longer than the wave length. (The effect of a finite extent of the wave crests has been discussed elsewhere[3].) The xy plane may coincide with the position of the undisturbed interface. The effect of the earth's rotation can be neglected because of the small scale of the phenomenon. Let U and ρ denote the wind and density of the lower layer, U' and ρ' the same quantities for the upper layer. As long as the wave amplitudes at the interface are small, the disturbance ζ of the interface* is given by

$$\zeta = Z \cos k\,(x - ct) \tag{2}$$

where Z is the constant amplitude, $k = 2\pi/L$, the orbital wave number, L the wave length, c the wave velocity. Further, for the velocity potentials in the lower and upper fluid

$$\phi = -Ux + (c-U)\,Ze^{kz} \sin k\,(x-ct) \tag{3a}$$

$$\phi' = -U'x + (c-U')\,Ze^{-kz} \sin k\,(x-ct) \tag{3b}$$

These forms of the solution satisfy the kinematic boundary conditions at infinity and at the interface. From the condition that the pressure be continuous at the interface, the following expression is derived for the wave velocity, as also shown by Lamb[10]

$$\frac{g}{k}(\rho - \rho') = \rho\,(c-U)^2 + \rho'\,(c-U')^2 \tag{4}$$

or

$$c = \frac{U + \rho' U'}{\rho + \rho'} \pm \left\{ \frac{g}{k}\,\frac{\rho - \rho'}{\rho + \rho'} - \frac{\rho\rho'}{(\rho + \rho')^2}(U - U')^2 \right\}^{1/2} \tag{5}$$

The first term on the right-hand side of (5) represents the mean velocity weighted according

* We are using only the real parts of the complex expressions given by Lamb.

to the density of each layer. Because of the small density difference at atmospheric dicontinuities, this term is practically equal to the unweighted arithmetic mean. The first term under the radical represents the stabilizing effect of the stratification assuming, of course, that ρ is larger than ρ'. The second term represents the instability due to the wind shear.

If the winds in both layers are either in the same or in opposite directions, as assumed so far, one would suppose *a priori* that the wave crests will be oriented perpendicular to the wind, and that the wave system which originates at the interface will move with the mean of the wind speeds in both layers. If the winds in both layers make an arbitrary angle, one might similarly expect *a priori* that the wave crests are oriented normal to the vertical wind difference because the ground is far away so that only the relative velocity of one layer against the other is of importance. For similar reasons one can expect that the wave system moves with the vector mean of the two winds. This problem has received considerable theoretical attention, and such observational data as are available for tropospheric billow clouds confirm these *a priori* expectations[3].

It should be pointed out that the angle between the wave crests and the mean wind under these conditions can vary between 0° and 90°, depending on the vector difference between the winds in both layers. It may, therefore, not always be possible to deduce the wind speed from the observations of noctilucent clouds, as has also been pointed out by Grishin[11].

When the wave velocity c equals the mean of the wind velocities in both layers

$$c=\frac{\rho U+\rho' U'}{\rho+\rho'}\cong\frac{U+U'}{2} \tag{6}$$

and it is seen from (5) that a double root exists in the equation for c. Hence the amplitudes of ϕ etc. are linear functions of time, and the waves with L given by (7), below, are those at the limit between stability and instability.

3. THE LENGTH OF THE WAVES

With the expression (6) for the wave velocity, one finds from (5) that the wave length

$$L=\frac{2\pi}{g}\frac{\rho\rho'}{\rho+\rho'}\frac{(U-U')^2}{\rho-\rho'} \tag{7}$$

The equation of state for ideal gases can be applied to the atmosphere so that

$$\rho=mp/RT \tag{8}$$

where p is the pressure, m the molecular weight, R the universal gas constant. The pressure must be continuous at the interface. Therefore (7) may also be written as

$$L=\frac{2\pi}{g}\frac{TT'}{T+T'}\frac{(U-U')^2}{T'-T} \tag{9}$$

The temperature difference is in general sufficiently small at atmospheric interfaces so that approximately

$$L=\frac{\pi}{g}T_m\frac{(U-U')^2}{T'-T} \tag{10}$$

where T_m is the mean of the temperature at the interface.

From the preceding three formulae it would appear that in the absence of a density and temperature discontinuity, the wave length should be infinite. But if the compressibility of the atmosphere is taken into account (10) must be replaced by the following expression[2]

$$L=\frac{\pi}{g}T_m\frac{(U-U')^2}{[(T'-T)^2+(U-U')^2T_m(\bar{\varepsilon}-\varepsilon)/g]^{\frac{1}{2}}} \tag{11}$$

where ε is the actual lapse rate of temperature assumed to be the same in both layers, $\bar{\varepsilon}$ ($\approx 1°K/100m$) the adiabatic lapse rate. The additional term in the denominator of (11) expresses the fact that the atmosphere is stable not only because of the density discontinuity at the interface, but also because the actual lapse rate is less than the adiabatic.

No data are yet available for noctilucent clouds to apply any of these expressions for the computation of the wave length. But it is of some interest to consider the orders of magnitude of the various parameters which would give the often observed wave lengths of about 10 km. Because of the observed altitudes of noctilucent clouds, it can be hypothesized that the wave patterns observed in them are associ-

ated with internal waves in the vicinity of the mesopause. If the temperature discontinuity is put equal to zero, it follows from (11) that

$$|U-U'|=\frac{L}{\pi}\left[\frac{g(\varepsilon-\varepsilon)}{T_m}\right]^{1/2} \qquad (12)$$

With $T_m=200°\mathrm{K}$, $g=9{\cdot}8\ \mathrm{m\ sec^{-2}}$, $\varepsilon=0$, $L=10{,}000$ m, one obtains $|U-U'|=70\ \mathrm{m\ sec^{-1}}$ which does not seem excessive in view of the large wind shears deduced for these layers from meteor observations. The temperature lapse rate in the vicinity of the mesopause is actually not zero; instead it undergoes a rapid change from positive to negative values so that the value of $U-U'$ is only an estimate. But in any case, no excessive value is required.

Abdullah[12] has extended the theory of billow clouds to wave groups. He finds that such groups of waves, "billow groups", should have horizontal dimensions about 10 to 100 times larger than the distances between the individual billows. This result is of interest in connection with the observation[11,13] that sometimes wave lengths of the order of 50 to 100 km have been observed at the same time as the shorter waves.

Witt[14] has recently reported an observation of noctilucent clouds where the cloud material moved in a southwesterly direction with a speed of 100 m $\mathrm{sec^{-1}}$ while a system of bright waves moved more slowly in the opposite direction. In view of the discussion in Section 4 below, one might interpret the bright waves as those regions where more cloud particles are concentrated per unit area because of larger horizontal convergence. Since the wave velocity c appeared to be opposite to the velocity of the cloud material U, the velocity U' in the upper layer would have to be opposite to U and its absolute value should be larger than U. This interpretation would lead to a value of more than 200 m $\mathrm{sec^{-1}}$ for the wind shear which appears to be larger than anything reported previously and would, with the other parameters chosen as above, imply a three times greater wave length. Unfortunately, no figure is available for the length of these waves so that it cannot be decided to what extent this observation is quantitatively in agreement with the suggestion made here.

4. FORMATION OF WAVE STRUCTURE IN DUST CLOUDS

It will now be shown that the horizontal divergence associated with the internal wave motion may well be large enough to produce observable billows. To do so we consider the variation in the horizontal divergence caused by the wave motion and its effect on the distribution of the particle density n. Let u, w be the horizontal and vertical velocity components due to the wave motion. Then we have for the local change of n

$$\frac{\partial n}{\partial t}=-\frac{\partial}{\partial x}[n(U+u)]-\frac{\partial (nw)}{\partial z} \qquad (13)$$

For the visual appearance of the cloud from the ground, the number of particles along the line of sight, rather than the particle density, is important. If the cloud is directly or nearly directly overhead, the particle number along the line of sight will be proportional to that above the unit area. This number will be determined first, and the effect of the low elevation angle of the noctilucent clouds will be discussed later.

It will be assumed that the dust which makes up the cloud is concentrated in a layer of thickness h below the interface, because the upper layer with its temperature increase with elevation has a much more stable stratification than the upper part of the mesosphere where the temperature decreases. Admittedly this argument for fixing the position of the noctilucent cloud layers is weak. However, for the explanation of the appearance of waves to be given here, it is irrelevant whether the cloud is below or above the interface so that this point need not be pursued further here.

To obtain an equation for the number N of dust particles above a $\mathrm{cm^2}$ we integrate from the bottom of the cloud, $-h$, to its top at 0. The variation in the position of the boundaries of the cloud, which is neglected, will only produce an error of higher order. If, for the sake of simplicity and because of lack of information, n is assumed independent of z within the cloud

$$\int_{-h}^{0} n\,dz=nh=N$$

By vertical integration of (13) through the cloud

$$\frac{\partial N}{\partial t}+U\frac{\partial N}{\partial x}=-\bar{n}\int_{-h}^{0}\frac{\partial u}{\partial x}dz \qquad (14)$$

On the right-hand side of this equation n has been replaced by its mean value $\bar{n}$ in horizontal direction, a substitution which is consistent with the other approximations. The last term on the right-hand side of (13) vanishes upon integration since it does not affect the number of particles above the unit area.

Let

$$N=N_0+N_1\cos k(x-ct)$$

where $N_0=\bar{n}h$, the mean value of N. Then, when $u=-\partial\phi/\partial x$ from (3a) is substituted in (14)

$$N_1=N_0\frac{Z}{h}(1-e^{-kh}) \qquad (15)$$

For the lengths of the waves, magnitudes of about 10 km have been frequently reported, and this value will be assumed here. For the thickness h of the cloud, Ludlam[5] has given an estimate based on an expression derived by Scorer[15] under the assumption that the billows form in a statically unstable layer of thickness d. In such a layer the billows should have the length $L=2{\cdot}7\,d$. But in view of the discussion in Sections 2 and 3, it does not seem likely that the billows are formed in this manner rather than as internal waves. On the other hand, Rapp[16] has recently made a theoretical study of the vertical distribution of a cloud of extra-terrestrial material penetrating the atmosphere below 100 km elevation and has found that such a cloud would be largely concentrated in a layer of fairly uniform density about 5 km thick. Therefore, it seems justified to assume here $h=4$ km. Since the noctilucent clouds are quite tenuous, it is not likely that h could be considerably larger than this value. Also, since meteor data indicate considerable wind shear throughout this region of the atmosphere, it does not seem likely that a very deep cloud could persist very long.

Few observations are available of the amplitude Z of the oscillating interface. According to Witt[13], measurements of noctilucent cloud waves in one particular case have indicated that $Z=1$ km. It is likely that this value is not too large considering that the wave amplitudes in the case of the billow clouds of the lower atmosphere are large enough to produce condensation and evaporation. With these figures

$$N_1=N_0(1/4)\,0{\cdot}92=0{\cdot}23\,N_0$$

Thus the percentual variation of N would be almost 50 per cent.

If the thickness of the noctilucent clouds is much smaller than their wave length, it follows from (15) that approximately

$$N_1=N_0 2\pi\frac{Z}{L} \qquad (16)$$

With the same values for Z and L as before

$$N_1=0{\cdot}63\,N_0$$

In both cases the variation of the particle amount per unit area caused by the divergence of the wave motion would be quite appreciable and of the same magnitude as the effect of perspective.

The effect of the low elevation angle of the noctilucent clouds modifies these estimates. If the cloud crests and troughs are in the direction towards the observer, the line of sight, despite its small inclination to the horizontal, will pass, completely or nearly completely, through the cloud from the bottom to the top because the lengthwise extent of the crests and troughs is considerably larger than the vertical thickness of the clouds. Therefore, the ratio of the particle numbers along lines through regions of greatest or smallest horizontal divergence will be about the same as given by (15) and (16). On the other hand, when the crests and troughs are normal to the line of sight, the divergence effect may be very much reduced because of the low elevation angle of the clouds. Since the influence of the viewing angle on the appearance of lighter and darker bands in the clouds will, on the other hand, be greatest when the crests and troughs are normal to the line of sight and smallest when they are parallel to it, the two effects supplement each other. It can, therefore, be concluded that the appearance of wave

structure in the noctilucent clouds does not require the assumption that they consist of water substance.

The foregoing argument is not substantially changed if allowance is made for the fact that the extent of the wave crests is not infinite[3]. The expression for the horizontal divergence remains the same as before. The wave length of the billows with finite crests is reduced, compared to the case of infinite lateral extent, but the reduction is small, even if the lateral extent is only a few times larger than the wave length.

For waves of much greater length, of the order of 50–100 km, as have occasionally been observed[11, 13], formula (16) gives a negligible variation of the horizontal divergence so that for long waves the explanation given here for the appearance of the billows would appear to break down. But a preliminary calculation allowing for the effect of the compressibility indicates that even for such long waves, the variation of the horizontal divergence may be quite appreciable. Moreover, it may well be that the shape of the waves deviates sufficiently from the sinusoidal form assumed here so that the variation of the horizontal divergence is sufficiently pronounced to produce the observed changes in the luminosity of even long cloud waves. It would be highly desirable to have calibrated densitometer measurements of noctilucent clouds as a basis for more thorough quantitative studies of this question.

5. CONCLUSION

The preceding discussion makes it appear likely that the billows observed in noctilucent clouds are manifestations of internal waves. A definite proof of the wave nature of these billows requires data of the wind and temperature distribution in the region where and at the time when noctilucent clouds are observed. Once the dynamics of the structure observed in these clouds is better understood, their further study will give important clues to the motions of the atmosphere in the vicinity of the mesopause.

REFERENCES

1. G. Witt, *Tellus*, **9**, 363 (1957).
2. B. Haurwitz, *Beitr. Geophys.* **34**, 213 (1931) and **37**, 16 (1932).
3. B. Haurwitz, *Ann. N. Y. Acad. Sci.* **48**, 727 (1947).
4. D. Deirmendjian and E. H. Vestine, *Planet. Space Sci.* **1**, 146 (1959).
5. F. H. Ludlam, *Tellus* **9**, 341 (1957).
6. I. A. Khostikov, Some Problems of Meteorology, *Com. Ann. Geophys. Int.* Acad. Sci. U.S.S.R. No. 1, 112 (1960).
7. V. A. Bronshten, Some Problems of Meteorology *Com. Ann. Geophys. Int.* Acad. Sci. U.S.S.R. No. 1, 120 (1960).
8. E. H. Vestine, *Journ. Roy. Astr. Soc. Can.* **28**, 249 (1934).
9. Zd. Sekera, *Beitr. Geophys.* **54**, 9 (1938).
10. H. Lamb, *Hydrodynamics* 6 ed., Chap. IX, art. 232, Dover Publ. N.Y. (1945).
11. N. I. Grishin, Some Problems of Meteorology, *Com. Ann. Geophys. Int.* Acad. Sci. U.S.S.R. No. 1, 117 (1960).
12. A. J. Abdullah, *Ann. N. Y. Acad. Sci.* **48**, 744 (1947).
13. G. Witt, Private Communication.
14. G. Witt, *Nature, Lond.* **184**, 1618 (1959).
15. R. S. Scorer, *Quart. J. R. Met. Soc.* **77**, 235 (1951).
16. R. R. Rapp, Verbal communication.

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 99–104. Printed in Great Britain.

THE SYSTEMATIC BEHAVIOR OF HYDROGEN EMISSION IN THE AURORA—II

G. C. REID* and M. H. REES

Geophysical Institute, University of Alaska, College, Alaska

(*Received* 10 *June* 1960)

Abstract—Trapping of solar corpuscular radiation by the geomagnetic field is examined in a semi-quantitative way. A mechanism is proposed which, within certain simplifying assumptions, explains the nocturnal latitude drift of the aurorally associated hydrogen emission described in the previous paper. It is shown that if the extent of this latitude drift is used to define the effective region of auroral occurrence at any time, then this region must move towards the equator and expand under highly disturbed conditions.

1. INTRODUCTION

Many years ago, Chapman and Ferraro[1] proposed a theory of the initial phase of a geomagnetic storm based on the effects produced by the earth's magnetic field on a stream of ionized gas of solar origin. Their chief conclusion was that the motion of the ionized gas in the geomagnetic field would induce currents in the advancing plane face of the stream whose magnetic effects could be represented by a magnetic dipole situated as far behind the face of the stream as the earth-dipole is in front of the stream. The moments and orientations of the two dipoles will be the same, assuming the stream to be perfectly conducting, and this second imaginary dipole will be referred to here as the "image" dipole. The configuration of the earth's magnetic field under these conditions will be approximately of the form shown in Fig. 1, the lines of force being effectively compressed on the day-side of the earth in the neighborhood of the approaching stream, and the field intensity being doubled at points in the equatorial plane very close to the stream front. The stream will continue to advance towards the earth with diminishing velocity until it is finally brought to a halt at such a distance from the earth that the energy density stored in the compressed geomagnetic field is roughly equal to the kinetic energy density of the particles in the stream. The stream-front at great distances from the earth travels on unimpeded, so that the geomagnetic field carves out a hollow in the stream-front, and the night side of the earth, away from the stream, is comparatively sheltered from the direct effect of the stream, at least during the early part of the magnetic storm.

More recently, Biermann[2] has made use of observations of comets to postulate the existence of a continuous outflow of matter from the sun in all directions. This suggestion has been developed by Parker[3], and the phenomenon has been given the descriptive name of "solar wind". This concept is very similar to that of Chapman and Ferraro, the only essential difference being that the latter

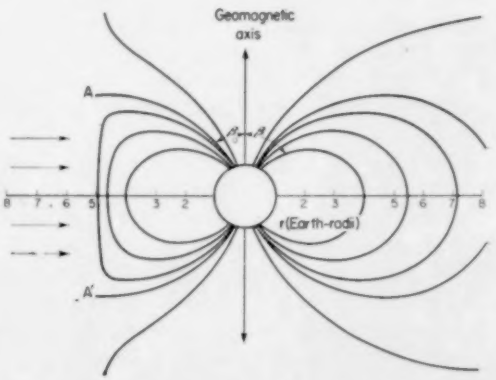


Fig. 1. The distortion of the geomagnetic field on the day-side of the earth due to an approaching solar particle stream.

* Present address: Defence Research Telecommunications Establishment, Shirley Bay, Ottawa.

authors treat the occurrence of particle streams as individual events separated by periods of complete calm, whereas Biermann and Parker postulate a continuous wind, which may increase in strength from time to time, with consequent increases in geomagnetic disturbance.

Another factor which must now be considered in any discussion of conditions in the neighborhood of the earth is the recent discovery by Van Allen *et al.*[4] of two vast belts of high-energy particles apparently trapped in the geomagnetic field.

The reason for the existence of two belts is not yet completely clear, but present indications are that the inner belt consists of protons resulting from the decay of the cosmic-ray neutron albedo[5]. The outer belt, which appears to consist primarily of electrons, has been tentatively ascribed to particles of solar origin[6].

The particles associated with aurora are probably of considerably lower energy than those which have been detected in the Van Allen belts, but the existence of the belts strongly suggests the possibility that trapping in the geomagnetic field can provide a storage mechanism for these auroral particles as well. In particular, the azimuthal drift of trapped particles in a dipole field provides a mechanism for transporting the particles from the sunward side of the earth, where they must initially appear, to the dark side, where the visual auroral features appear. In this connection, it might be noted that Dessler and Parker[7] have recently suggested that the temporary trapping of protons of solar origin plays a dominant role in the geomagnetic storm effects which are closely associated with the occurrence of visual aurora.

Throughout the remainder of this paper, the temporary trapping of solar plasma will be accepted as a working hypothesis. The trapping will be assumed to occur as a result of the deformation of the geomagnetic field produced by the solar plasma cloud, and the positively charged component of the plasma will be assumed to consist of protons. The purpose of the present paper is to show how the observed diurnal variation in latitude of the hydrogen emission associated with aurora[8] can be tentatively explained on the basis of an interaction between the solar wind, the geomagnetic field and the trapped protons.

2. DRIFT MOTIONS OF THE TRAPPED PROTONS

The individual trapped particles perform complicated orbits, the principal component of which is a spiralling motion about the direction of the magnetic field accompanied by a drift along the direction of the field. This drift is reversed in direction at two points, one in each hemisphere, whose position is determined by the injection pitch angle of the spiral. Superimposed on this motion is a much slower azimuthal drift around the earth caused by the spatial inhomogeneity of the earth's magnetic field, and in the direction of the vector product of the field and the field gradient. Since this latter drift is at right angles to the direction of the gradient of the field, the individual particles of low pitch angle tend to remain in a region of constant magnetic field strength as they circle the earth close to the equatorial plane. If the earth were a perfect dipole, this would result in a strictly longitudinal drift in the equatorial plane. Some of the possible consequences of the departure of the geomagnetic field from that of a dipole have been discussed by Gold[6], and will not be mentioned further here. We shall assume that the dipole approximation is valid, and examine the consequences of postulating a solar wind of the type discussed above. We shall make the further simplifying assumption that the solar wind direction is normal to the direction of the earth-dipole, so that the deformation of the field is as shown in Fig. 1. This is not strictly true, but the approximation is probably adequate for the semi-quantitative discussion given below.

Inspection of Fig. 1 shows that there are two points, A and A′, at which the field lines are orthogonal to the stream-front and at which the resultant field vanishes. In their original discussion, Chapman and Ferraro[1] pointed out that "horns" of ionized gas might protrude towards the earth from these points, but they did not consider these horns to be of any great significance in connection with the aurora, since

they would exist only on the day-side of the earth. In a more recent article, Warwick[9] has again directed attention to these regions as a source of intruding material. In this discussion, however, we shall attach a somewhat different significance to them, and say merely that the lines of force coming down from these points intersect the earth at some point which represents the upper limit of the region affected by the trapped protons. Beyond these points, trapping of particles would not be possible. Here again we are forced to over-simplify the picture, since the positions of these zero-field points will not remain the same as the stream sweeps past the earth. Their position will depend on the configuration of the "wind-break" caused by the earth, which is by no means determined as yet. For simplicity we shall once again use the situation depicted in Fig. 1, and we shall also assume that the field lines connecting A and A′ to the earth are those of the undisturbed earth-dipole. Under these conditions it can readily be shown that the geomagnetic co-latitude β_0 of the points of intersection with the earth are given by

$$\sin^2 \beta_0 = (1{\cdot}5)^{-3/2} a/d \qquad (1)$$

where a is the radius of the earth and d is the distance from the center of the earth to the stream-front. Equation (1) can be re-written

$$\sin^2 \beta_0 = 0{\cdot}54/d \qquad (2)$$

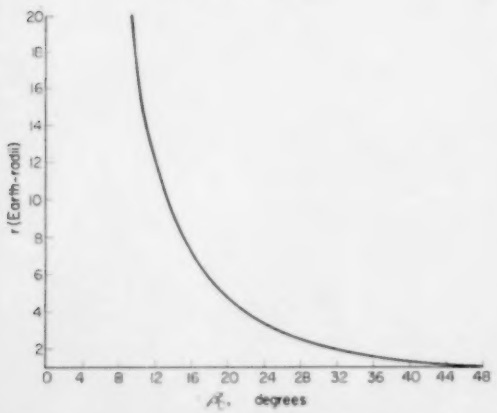


Fig. 2. Geomagnetic co-latitude (β_0) at which the lines of force from A and A′ intersect the earth's surface as a function of distance of the stream-front (d) from the center of the earth.

where d is now measured in earth-radii. This relation is shown in Fig. 2.

The field lines from regions immediately to the south of these points (in the northern hemisphere) might then be expected to correspond to the outer edge of the trapped proton belt, and along the line of force which touches the stream-front, the field intensity in the equatorial plane would be twice as great as its undisturbed value.

Let us now consider the actual distribution of field intensity as a function of radial distance from the earth along a line drawn in the equatorial plane normal to the stream-front. The field is that due to two equal dipoles separated by a distance $2d$, and is given by

$$H = \frac{M}{|r|^3} + \frac{M}{|2d - r|^3} \qquad (3)$$

where M is the magnetic moment of either dipole ($=8{\cdot}06 \times 10^{25}$ cgs units) and r is the distance from the center of the earth. r is positive on the day-side of the earth (towards the stream-front) and negative on the night-side. On the night-side of the line, the second term becomes much smaller than the first, and the field can be represented to a good approximation by the undisturbed dipole field of the earth. This is plotted in the lower curve of Fig. 3. The distorted day-side field is shown in the upper curve, assuming the stream-front to be situated at a distance of 5 earth-radii. The two field distributions are seen to be closely similar near the earth and to deviate more and more as we approach the stream-front. At the stream-front the day-side field is twice the night-side field at the same distance from the earth.

Returning to Fig. 2, we see that for a stream-front distance of 5 earth-radii, the geomagnetic co-latitude, β_0 corresponding to the termination of the trapped proton belt is about 19°. Fig. 3 shows that particles which are trapped along the corresponding line of force experience a local field of 5×10^{-3} gauss in the equatorial plane on the day-side. As they drift around the earth in a region of constant field strength, they will appear at a distance of 4 earth-radii at magnetic midnight (see Fig. 3). Since on the night-side of the earth, the simple dipole field is a good

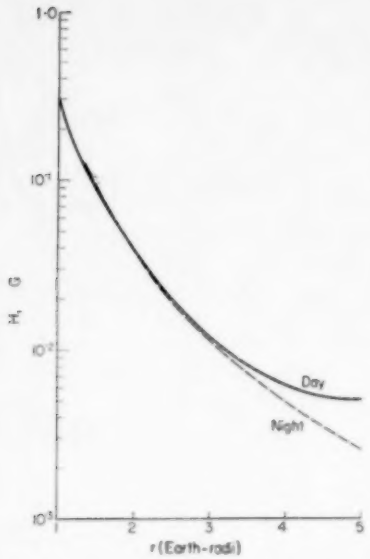


Fig. 3. Field strength in the plane of the geomagnetic equator for the conditions illustrated in Fig. 1. The stream front is assumed to be stationary at a distance of 5 earth-radii from the center of the earth on the day-side.

approximation, the corresponding geomagnetic co-latitude at which the line of force intersects the earth is given by

$$\sin^2\beta = 1/r = 0{\cdot}25 \qquad (4)$$

leading to a value for β of 30°. Thus we see that as the earth turns under this stationary pattern, the geomagnetic latitude at which the influence of these particles is felt decreases from 71° during the day to 60° at magnetic midnight, then drifts north again. This is precisely the behavior of the hydrogen emissions associated with aurora, illustrated by a six day sequence in Fig. 4.

An estimate of the distance d at which the stream-front will halt can be obtained by equating the kinetic energy density of the stream to the increased magnetic energy density of the compressed field. Since the local field strength is doubled at the surface of the stream, this can be written as

$$\tfrac{1}{2}\rho v^2 = (1/8\pi)\,H^2 \qquad (5)$$

where ρ is the density of matter in the stream, v is the stream velocity and H is the undisturbed value of the geomagnetic field at distance d from the center of the earth in the equatorial plane. This can be written

$$d = \left[\frac{H_0^2}{4\pi\rho v^2}\right]^{1/6} \qquad (6)$$

where d is measured in earth-radii and H_0 is the surface value of the field strength at the equator.

Taking $H_0 = 0{\cdot}35$ gauss and $v = 1000$ km/sec, we find that the value of 5 earth-radii for d used in the example above corresponds to a stream density of $6{\cdot}24 \times 10^{-23}$ gm/cm³. If we assume that the stream consists mainly of protons, their number density must then be about 40 per cm³. If the velocity is taken as 500 km/sec, the number density would have to be about 160 per

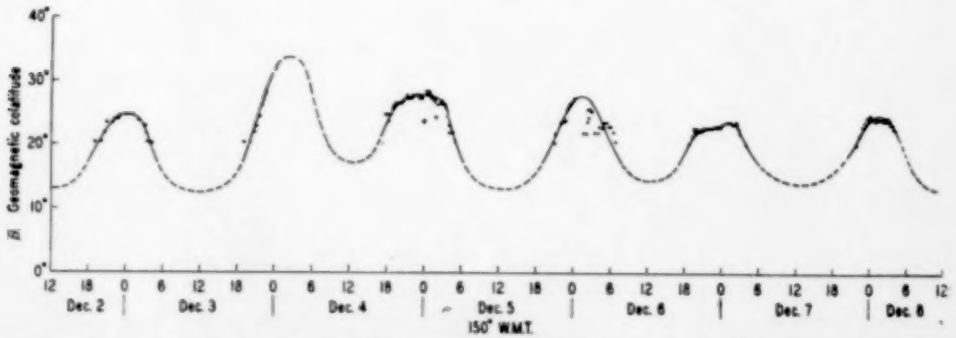


Fig. 4. Observational results of the latitude drift of auroral hydrogen emission for a six day period. The dashed portion of the curve is extrapolated.

cm³. Estimates of the solar wind strength[9] indicate that these values probably represent fairly normal conditions.

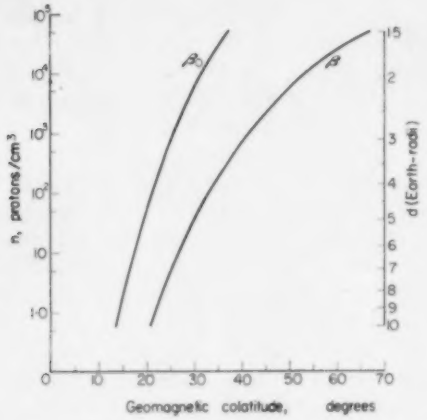


Fig. 5. Geomagnetic co-latitude of points on the earth's surface connected magnetically to the outer edge of the trapped proton belt on the side towards the solar wind (β_0) and away from the solar wind (β). The right-hand scale shows the stationary position of the stream-front and the left-hand scale gives the corresponding solar wind density, assuming the particles to be protons travelling with a velocity of 1000 km/sec.

Fig. 5 shows a plot of the co-latitude β_0 defined above, and also the corresponding midnight co-latitude β as functions of solar wind strength in protons per cm³, assuming a wind velocity of 1000 km/sec. The corresponding values of d, the distance of the stationary stream-front from the center of the earth, are also shown. The curves were obtained by using the method outlined in the example given above for the particular case of $n=40$ protons per cm³, $v=1000$ km/sec, $d=5$ earth-radii. These curves can be used to estimate the southward drift in latitude of the region on the earth at the end of the field line corresponding to the outer limit of the trapped proton belt. For instance, for $n=10^4$ protons per cm³, this region will drift from 58° to 37° geomagnetic latitude between magnetic noon and midnight. For $n=1$ proton per cm³, the corresponding values are 76° and 68°. Comparison with the observed latitude drifts of auroral hydrogen emissions shown in Fig. 4 reveals a remarkable similarity between the observed and predicted amounts of latitude drift. Although nothing definite can be said about the daytime conditions, the observed night time points indicate that a total drift of ten or fifteen degrees is common. On Fig. 5 this is seen to correspond to moderately strong solar winds, with the outer edge of the trapped particle region at from 3 to 6 earth-radii.

The examples used above to illustrate the latitude drift also predict a finite range of latitudes over which hydrogen emission might be observed for a given proton density. It is interesting to note that the separation between the two curves of Fig. 5 for low particle densities is of the order of 5 or 6 degrees, or about the width of the classical auroral zone. If we carry this analogy to its logical conclusion, Fig. 5 can be taken as predicting that the "auroral zone" simultaneously moves towards the equator and expands during periods of high density solar corpuscular radiation.

3. THE HYDROGEN EMISSION

So far little has been said concerning the mechanism of production of the hydrogen emissions, and we have merely implied that they are connected with the trapped protons. This assumption has been made on the basis of the facts that (1) protons would be expected to emit the observed radiation as they recombine with electrons; (2) the azimuthal drift of the trapped protons provides a simple means of transferring them from the day-side of the earth to the night-side, where the auroral effects are observed. The approximate coincidence between the predicted and observed latitude drifts discussed above lends further strength to this assumption, and also helps to justify the Chapman–Ferraro picture of the magnetic effects of a solar particle stream in the neighborhood of the earth.

The direct cause of the hydrogen emission can be tentatively attributed to the removal of protons of large pitch-angle from the trapping region, accompanied by the emission of light as they recombine with electrons in the ionosphere. The removal may be largely due to charge exchange with the neutral hydrogen of the earth's exosphere, as suggested by Stuart[11], with some contribution from large angle elastic

scattering. The latitude drift which has been estimated above will cause a lowering in the mirror altitude of particles of large pitch-angle, even if the earth's magnetic field is assumed to be that of a dipole at the geometrical center of the earth. This is due to the decrease of field strength with decreasing latitude at constant distance from the dipole. However, in any quantitive discussion, the effects of higher-order terms in the magnetic field distribution will probably prove to be very much greater. These effects have been discussed qualitatively in a recent article by Gold[6].

4. COMMENTS

Many approximations have been made in the above discussion for the sake of simplicity and ease of calculation. Chief among these is the neglect of the so-called "integral invariant" of trapped particle motion[10] $\int v_{\parallel} dl$, which should be conserved throughout the longitudinal drift. The effect of applying this invariance law is reserved for a future discussion.

Among other simplifications, the outer edge of the trapped proton belt, corresponding to the solar stream-front, has been taken to be sharply defined. Obviously this cannot be the case, since the induced currents which give rise to the disturbing magnetic field must be set up in a finite volume of the gas. However, the approximation is probably valid in a semi-quantitative discussion of this type. Also no account has been taken of the effects of the reduced radial field gradient on the day-side of the earth due to the compressing effect of the solar stream. This will have the effect of reducing the azimuthal drift velocity of particles in the equatorial plane. The contribution of trapped electrons to the auroral effects has also been ignored, although it seems highly probable that this would have to be included in any complete investigation, since the original solar plasma must presumably be electrically neutral. The trapped protons and electrons will drift around the earth in opposite directions, and if the position of the solar stream-front is changing with time, the possibility arises of protons and electrons appearing simultaneously at different latitudes on some night-side meridian of the earth. This effect may be responsible for temporary charge separation in the belt, which would in turn give rise to electric fields. These fields, when transferred to the ionosphere via the magnetic field lines may be responsible for the more active auroral forms which follow the appearance of the hydrogen emissions. This has been discussed by one of us in an earlier publication[12].

Until more is known, both on theoretical and experimental grounds, about the behavior of trapped particles in a distorted dipole field, all hypotheses of this type must be purely speculative, but the limited success of the above calculations in explaining one feature of the aurora lends some strength to the suggestion that the trapping of particles in the geomagnetic field plays an important part in the auroral phenomenon.

REFERENCES

1. S. Chapman and V. C. A. Ferraro, *Terr. Magn. Atmos. Elect.* **36,** 77, 171 (1931); **37,** 147. 421 (1932).
2. L. Biermann, *Z. Astrophys.* **29,** 274 (1951); *Z. Naturf.* **7a,** 127 (1952); *Obervatory* **77,** 109 (1957).
3. E. N. Parker, *Phys. Fluids* **1,** 171 (1958).
4. J. A. Van Allen, and L. A. Frank, *Nature, Lond.* **183,** 430 (1959).
5. S. F. Singer *Phys. Rev. Letters* **1,** 171 (1958).
6. T. Gold, *Nature, Lond.* **183,** 355 (1959).
7. A. J. Dessler and E. N. Parker, *J. Geophys. Res.* **64,** 2239 (1959).
8. M. H. Rees, G. J. Romick and A. Belon, *Planet. Space. Sci.* (in press).
9. J. W. Warwick, *J. Geophys. Res.* **64,** 389 (1959).
10. M. N. Rosenbluth and C. L. Longmire, *Ann. of Phys.* **1,** 120 (1957).
11. G. W. Stuart, *Phys. Rev. Letters* **2,** 417 (1959).
12. G. C. Reid, *Nature, Lond.* **182,** 1791 (1958).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 105–108. Printed in Great Britain.

INVESTIGATION OF AURORAL ECHOES—II

L. HARANG and J. TRÖIM
Norwegian Defence Research Establishment, Kjeller, Norway

(*Received* 20 *May* 1960)

Abstract—In part I of this paper the results of angle of arrival and drift of the auroral ionization have been discussed.

In the present paper the *reflection coefficient* of the echoes on 40, 80 and 120 Mc/s will be discussed.

The frequency dependence of the reflection coefficient has been studied. For identical transmitting and receiving equipment the following representative values were found. An echo received on 40 Mc/s with an amplitude of 10 μV was accompanied by a similar echo on 80 Mc/s of 3 μV amplitude. An echo received on 40 Mc/s of 20 μV was usually accompanied by an echo on 120 Mc/s of the order of amplitude of 1·5 μV.

Various models of scattering mechanisms are discussed, and the most probable one is considered to be field-aligned with irregularities arranged in large surfaces in E–W direction.

1. INTRODUCTION

Observations have shown that there is a frequency and range dependence of the echo amplitudes received, the amplitudes decrease strongly with increasing frequency and range. Investigations have been made on the frequency- and range-dependence in the VHF-band 40–100 Mc/s, but a generally accepted law has not yet been established.

As a limiting ideal case we may consider two alternatives: scattering from a single target with equivalent echoing area (or cross-section) σ, and partial reflection from an infinite plane surface with reflection coefficients:

$$\sigma = A^2 \frac{P_r' 64\pi^3}{P_t G^2 \lambda^2} R^4 \quad \text{and} \quad \rho = A^2 \frac{P_r' 32\pi^2}{P_t G^2 \lambda^2} (2R)^2 \tag{1.1}$$

where

- P_r' is the power received corresponding to an amplitude of 1μV
- A echo amplitude in μV
- P_t power transmitted
- R range
- G power gain of the identical receiving and transmitting aerials
- λ wavelength

In a discussion of observations from New Zealand, Seed[1] states that the observations of the echo amplitude versus range indicate an R^{-2} rather than an R^{-4} law. When measuring the echo amplitudes at different ranges the amplitudes have to be corrected for variations in aerial gain with elevation, which at these low angles of elevation (6°–15°) will be quite appreciable, but according to Seed it can be asserted that the range exponent is not four but more likely to be two, or to have a slightly higher value.

In Part I of this paper[2] the studies of drifts of the ionized areas using two aerials with divergence in azimuth or an interferometer with a number of lobes in the horizontal plane, clearly indicate that in most cases the ionized areas will have a considerable E–W extension which will completely fill out the opening of the aerial in E–W direction. Only in a few cases the interferometer records indicated the presence of limited ionized areas which could be approximated to single targets of scattering. A scattering area which partly fills out the opening of the aerials will favour an R^{-2} dependence of amplitudes, and the range observations stated by Seed seem to be in accordance with this point of view.

Concerning the wavelength dependence of σ or ρ all observations indicate that the amplitude, A, received decreases with decreasing wavelength. If the decrease observed was according to $A \propto l/\lambda$, σ or ρ would appear as constants

and independent of wavelength for the frequency range considered. The observations discussed in this paper show, however, that A decreases more rapidly than $1/\lambda$, which indicates that ρ or σ also depends on the wavelength.

2. EXPERIMENTAL ARRANGEMENT

The observations were made at Tromsö ($\phi = 70°$) lying about 300 km to the south of the normal auroral zone. All observations were made during winter time. The three-channel echo transmitter with channels on 40, 80 and 120 Mc/s has been described in Part I (2). Three identical eight elements Yagis with the elements horizontal were erected on the plane surface at the southern edge of a lake. The electrical heights of the aerials were the same on all three frequencies, $h = 1\lambda$, and having a well-defined reflection surface in front of the aerials considerably greater than the first Fresnel zone, identical aerial patterns for the three frequencies were assumed. From the free space aerial pattern and the assumed ground reflections it is possible to calculate the vertical aerial pattern. The maximum gain was $G = 20$. The maximum of the main lobe will cut the E-layer height between 300 and 600 km range, and the amplitudes measured were therefore restricted to echoes appearing within this range interval. The aerials were supplied with R–T arrangements, the same aerial being used both for transmission and reception. The peak powers of the transmitter on 40, 80 and 120 Mc/s were respectively 20, 15 and 10 kW, and the pulse length 100 μs. Common pulse modulators and HT-supplies were used for all three frequencies through which individual variations in pulse lengths and powers were eliminated.

The input of the receiver was matched to the aerial and calibrated by a signal generator. For simultaneous reception of the echoes on each of two frequencies (40 and 80 Mc/s or 40 and 120 Mc/s) two separate receivers were used and the echo patterns were displayed on a two-beam CRT.

In order to record the large variations in amplitude and using a range–time display from an intensity modulated CRT on continuous photographic paper, a cyclic variation of the receiver sensitivity was introduced. The IF-gain was varied through a motor-driven, externally connected 360° ranged potentsiometer. The motor speed was 1 r.p.m. and the gain control showed a logarithmic variation. This cyclic variation of the receiver threshold sensitivity was calibrated by a pulsed signal generator. On 40 Mc/s the attentuator was thus adjusted for outputs of 0·1 μV, 1 μV, 10 μV and 100 μV during successive cycles. The length of the blackened traces thus recorded gave a direct measure of the echo amplitude.

Dual recordings on 40 and 80 and on 40 and 120 Mc/s were made continuously, and a record is shown in Fig. 1. The range in threshold sensitivity could be set independently for the three frequencies. Usually it was adjusted as follows:

40 Mc/s variation in threshold sensitivity 0·1 to 100 μV
80 Mc/s variation in threshold sensitivity 0·1 to 20 μV
120 Mc/s variation in threshold sensitivity 0·1 to 5 μV

3. OBSERVATIONS

With the aerial arrangement used the echo patterns recorded on each pair of frequencies appeared at the same range interval, and the general character of the amplitude distribution within the echo-width was similar.

In Part I of this paper the question of aspect sensitivity has been discussed. It was shown that for 40 Mc/s echoes the condition for perpendicularity was almost fulfilled, the difference being only 2°–3° from normal incidence at ranges of 500–600 km, for greater and smaller ranges there could be a deviation from normal incidence of 6°–9°. If we assumed a slight bending of the 40 Mc/s ray in the Es-layer present it was further shown that assuming a moderate value of the electron density in the Es-layer, the effect of bending was sufficient for changing the path of the 40 Mc/s ray to be perpendicular on the geomagnetic field lines at the height of scattering. The records on Fig. 1 of the 120 Mc/s echoes, which show maxima at the same ranges as the 40 Mc/s echoes, indicate that the condition of perpendicularity need not

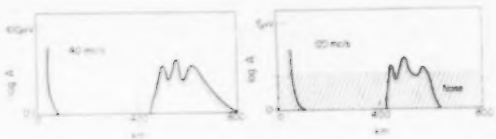


Fig. 1.

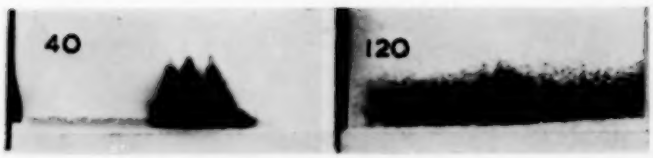


Fig. 1b.

Fig. 1. Simultaneous reception of auroral echoes on 40 and 120 Mc/s.

be strictly fulfilled. On a number of records on 80 Mc/s it is, however, apparent that there is a difference from the 40 Mc/s echoes in structure and this appeared at the part of short ranges up to 300 km. The amplitudes on the short range were very low or even lacking on some of the 80 Mc/s records. At these short ranges there will be a deviation from normal incidence of about 8°. The paths of the 40 Mc/s echoes had apparently been sufficiently bent in the *Es*-layer to approximate normal incidence, but this has not been possible for the paths of the 80 Mc/s echoes.

Auroral echoes will normally appear with considerable amplitudes over a certain range, the width is usually of the order of 150 km. In the estimates of amplitudes only the maximum of the amplitude value has been measured. If the echoes had been recorded as CW-reflections from the ionized areas, as in the experiments by Forsyth and Vogan[3] the echo amplitudes would appear integrated over the complete width of the echoes, and the absolute value of the reflections coefficient would therefore in this case be greater. For a determination of the wave length dependence of the reflection coefficient this ambiguity is without importance.

The results of the simultaneous recordings of echoes on 40, 80 and 120 Mc/s were as follows. An echo amplitude of $A_{40}=10$ μV on 40 Mc/s was usually accompanied by an echo on 80 Mc/s of a mean value of $A_{80}=3$ μV (after correction being made for the difference in transmitting power). An echo amplitude of $A_{40}=20$ μV on 40 Mc/s was usually accompanied by echoes on 120 Mc/s. The amplitude varied strongly from one case to another, usually lying in the interval 2 to 0·5 μV. An average value of $A_{120}=1{\cdot}5$ μV is considered to be a representative value.

4. DISCUSSION

The observations clearly indicate that the decrease in amplitude according to the law $A \propto 1/\lambda$ assumed in formula (1) is not correct, which again means that σ or ρ must depend on the wave length. Using formula (1) the following ratios are obtained, which are considered to be representative:

$$\frac{\sigma_{40}}{\sigma_{80}}=\frac{\rho_{40}}{\rho_{80}}=\left(\frac{A_{40}}{A_{80}}\right)^2\left(\frac{\lambda_{80}}{\lambda_{40}}\right)^2=2{\cdot}8$$

and

$$\frac{\sigma_{40}}{\sigma_{120}}=\frac{\rho_{40}}{\rho_{120}}=\left(\frac{A_{40}}{A_{120}}\right)^2\left(\frac{\lambda_{40}}{\lambda_{120}}\right)^2=20$$

Forsyth and Vogan[3] have studied the changes in reflection coefficient with wave length using three frequencies in the band 32·2–50·0 Mc/s. The amplitude ratios of the signals were compared with different models of scattering mechanisms each of which should involve different wave length dependence: scattering from nonisotropic irregularities, scattering from isotropic irregularities, partial reflections from large surfaces, scattering from long, thin columns and scattering from intensely ionized regions. A definite and single model of the scattering process could not be evaluated from the observations, but according to Forsyth and Vogan the measurements seemed to point to the existence of auroral scattering centres within which the electron density was of the order of 10^7 el/cm^3.

The appearance of echoes simultaneously and with similar echo structure on 40 and 120 Mc/s seems difficult to reconcile with the model of scattering centres of free electrons. The aspect sensitivity and the high coefficient of reflection which during great displays may appear on 40 Mc/s (input voltages up to 100 μV have been measured) seems to indicate that the process of back scatter is most probably a combination of scattering from long, thin columns arranged in surfaces with their reflecting planes lying in E–W direction.

As one limiting case we may consider the partial reflection process from large surfaces, and apply the measurements of reflection coefficient on 40, 80 and 120 Mc/s.

For echoes appearing simultaneously on 40, 80 and 120 Mc/s with the representative values of echo amplitudes $A_{40}=20$ μV, $A_{80}=3\times\sqrt{2}$ μV and $A_{120}=1{\cdot}5$ μV, and a range of 500 km, one obtains the following coefficient of reflections:

$\rho_{40}=5{\cdot}66\times10^{-6}$, $\rho_{80}=\rho_{40}/2{\cdot}8$, $\rho_{120}=\rho_{40}/20$.

For a partial reflecting surface the reflection coefficient (power) is given by:

$$\sqrt{\rho}=\frac{1-n}{1+n}=\frac{1-n^2}{(1+n)^2}\approx\tfrac{1}{4}(1-n^2)$$

Now the refractive index is given by:

$$n^2 = 1 - \frac{81 \times 10^{-6}}{f^2} N$$

where f is given in Mc/s and N in el/cm³. We thus obtain:

$$N = 4{\cdot}9 \times 10^{-4} f^2 \sqrt{\rho}$$

For typical cases of strong auroral echoes the following numerical values are obtained:

$N = 1{\cdot}9 \times 10^5$ el/cm³ for 40 Mc/s echoes
$= 4{\cdot}5 \times 10^5$ el/cm³ for 80 Mc/s echoes
$= 3{\cdot}8 \times 10^5$ el/cm³ for 120 Mc/s echoes

The values of echo amplitudes used represent cases with strong reflections. The formation of sheets of electrons with discontinuities of the order of $2-4 \times 10^5$ el/cm³ during these conditions does not seem unreasonable.

Another limiting case is to consider the scattering from nearly transparent elongated field-aligned irregularities discussed by Booker[4]. The most important frequency-dependent factor in Booker's formula for back scatter is the exponential

$$\exp\left(-\frac{8\pi^2 T^2}{\lambda^2}\right)$$

where T is the correlation distance transverse to the axis of symmetry of the irregularities.

Using the values measured of σ_{40}/σ_{80} and σ_{40}/σ_{120} it is possible, for a first approximation, to estimate the value of T. For the ratio $\sigma_{40}/\sigma_{80} = 2{\cdot}8$ we obtain $T = 0{\cdot}48$ m and for the ratio $\sigma_{40}/\sigma_{120} = 20$ we obtain $T = 0{\cdot}51$ m. Now, Booker's formula is not very sensitive to variations in the σ-ratio. Therefore the coincidence between the two values may not be significant.

From the discussion above it is apparent that the frequency dependence of ρ or σ seems to be qualitatively explained either according to a λ^2-law (partial reflection from large surfaces) or according to an $\exp(-k/\lambda^2)$-law (scattering from field-aligned irregularities).

Acknowledgments—We wish to express our thanks to the director of the Auroral Observatory, Mr. E. Tönsberg and also to the Staff of the Observatory for hospitality and assistance during the observations at Tromsö.

REFERENCES

1. T. J. SEED, *J. Geophys. Res.* **63,** 517 (1958).
2. L. HARANG and J. TRÖIM, NDRE-Report No. 183, see also paper in *Planet. Space Sci.* **5,** 33 (1961).
3. P. A. FORSYTH and E. L. VOGAN, *J. Atmos. Terr. Phys.* **10,** 215 (1957).
4. H. G. BOOKER, *J. Atmos. Terr. Phys.* **8,** 204 (1956).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 109–114. Printed in Great Britain.

THE HEIGHT VARIATION OF THE AMBIPOLAR DIFFUSION COEFFICIENT FOR METEOR TRAILS

J. S. GREENHOW and J. E. HALL
c/o Sir Bernard Lovell, Jodrell Bank

(*Received* 13 *May* 1960)

Abstract—The results are presented of an experiment relating the heights of radio echoes from meteors with the ambipolar diffusion coefficient determined from the rate of decay of the echo amplitude. The scatter of the results is explained in terms of the errors in the measurements and a technique of analysis is given, which takes account of these errors. The slope obtained for the height/ln D relation is compared with those given by other workers and reasons for the discrepancies are discussed. It is concluded that the atmospheric scale height during winter at a height of 98 km above Jodrell Bank (53° 14′ N) is 8·0 km.

1. INTRODUCTION

Measurements of the rate of diffusion of ionized meteor trails are of considerable interest since the diffusion coefficient is directly related to the atmospheric density and temperature. The correct interpretation of diffusion coefficient–height profiles should thus reveal information about the upper atmosphere.

It has been shown by Kaiser[1] that the effective diffusion coefficient, D, for the positive ion and electron column is twice the ion diffusion coefficient, D_i, derived from kinetic theory, where D_i is given by the expression

$$D_i = \frac{8}{3\pi n \sigma_i^2}\left(\frac{k\theta}{\pi\mu}\right)^{\frac{1}{2}} \text{cm}^2 \text{ sec}^{-1} \qquad (1)$$

n is the atmospheric particle density, θ the atmospheric temperature, σ_i the collision diameter for meteor ions with air molecules and μ is the mean mass of a meteor ion (assumed equal to the mass of an air molecule). For meteor trails of low electron density, Herlofson[2] has shown that the echo amplitude decays exponentially with a time constant $\tau = \lambda^2/16\pi^2 D$, so that D may be determined from the decay times of meteors.

For an isothermal atmosphere n is proportional to the atmospheric pressure, p, where p varies with height in the manner $p = p_0 \exp (h_0 - h)/H$; H is the atmospheric scale height and h the height above an arbitrary height h_0 at which the pressure is p_0. It therefore follows that the gradient of the height/ln D graph, H_D, should equal the atmospheric scale height

$$\frac{dh}{d \ln D} = H_D = H \qquad (2)$$

Measurements of the variation of D with height have been made by Greenhow and Neufeld[3] and by Murray[4] but the values of the gradient obtained by these workers are in serious disagreement. Greenhow and Neufeld give a slope of 6·5 km, while Murray gives a mean gradient of 26 km. The reason for this disagreement lies in the different treatment of the results. The purpose of this paper is to present the results of an experiment which is less susceptible to the uncertainties present in the analysis of diffusion coefficient–height measurements, and also to give a method of analysis for height/ln D plots which gives a better value of the gradient.

2. TECHNIQUE AND METHOD OF ANALYSIS

(a) By using a single transmitter and two widely spaced receiving stations, echoes have been obtained from different points on a single meteor trail. If the height separation of the two echoing points is known, it is possible to obtain the gradient of the height/ln D relation from the ratio in decay times of the two echoes from one meteor. As a height difference is measured, instead of a small difference between two almost equal heights, this method is not influenced by

height errors to the same extent as method (*b*). Also the ratio in diffusion coefficient is measured for a single meteor trail when any systematic errors in the estimation of decay time are likely to be the same, thus further reducing the effect of the large scatter in the absolute measurements of diffusion coefficient.

The equipment has been described by Greenhow and Neufeld[5] together with the method of determining the height separation of the echoing points. Only echoes showing an exponential decay have been used and τ has been estimated from an amplitude–time display by measuring the time taken for the echo to decay to half the initial amplitude. The observations were made at a frequency of 36 Mc/s.

(b) In the second experiment the actual meteor heights were found using the split beam technique described by Clegg and Davidson.[6] The echo decay times were again measured from an amplitude–time display and only meteors with exponential decays have been used. The corresponding values of height and ln D (diffusion coefficient) have been plotted on scatter diagrams. An example of the results is shown in Fig. 1, which gives the plot for meteors occurring between 03 hr and 06 hr U.T. during part of January and February 1959.

The errors in height of the meteors have been estimated from a calibration of the equipment and are about ± 4 km. It is more difficult to estimate the error in the decay time, but from considerations of the possible distortion of the exponential decay by resonance effects (Billam and Browne[7]), finite meteor velocity (Loewenthal[8]), the inclusion of meteors with line densities near 10^{12} electrons/cm. (Kaiser and Closs[9]) and the distortions due to upper atmospheric winds (Greenhow[10]) the error is estimated to be a factor of 1·3. Simultaneous measurements of decay time at two frequencies show that even echoes with apparently good

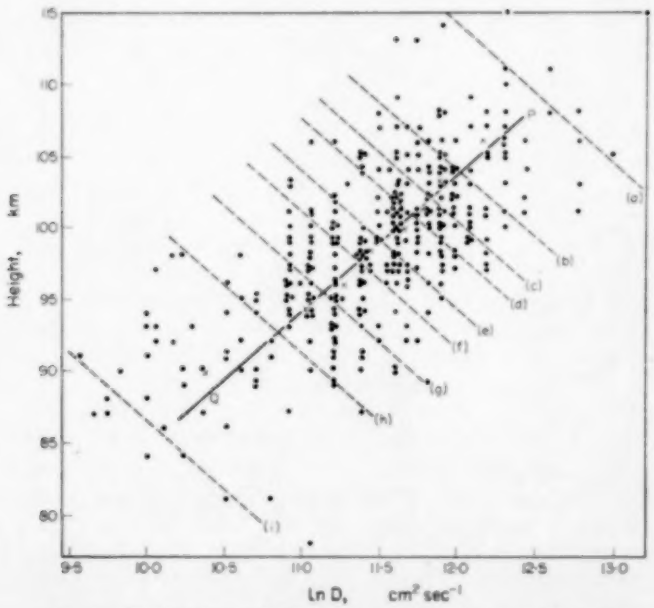


Fig. 1. The heights and diffusion coefficients for meteors occurring between 03hr–06hr U.T. during January–February 1959. (a) – (i) are lines dividing the points into groups (see text).
● Individual meteor echoes
× Median values for the groups
PQ Least-squares fit to the median points.

exponential decays can give diffusion coefficients differing by a factor of 1·6 (Greenhow and Hall[11]). This also indicates that the error at a single frequency is at least a factor of 1·3.

All the scatter of the individual points in Fig. 1 can be accounted for by the errors in D and h and in order to fit a line to the points it is necessary to take into account the magnitude of the errors in both co-ordinates. A method has been used which is equivalent to a least squares solution but, in order to reduce the labour, the following procedure has been used.† The points have been divided into eight groups by the lines (a)–(i) in Fig. 1 which have the gradient

$$-\frac{1}{H_D'}\left(\frac{\Delta h}{\Delta \ln D}\right)^2$$

where H_D' is the approximate gradient of the line of best fit and Δh, $\Delta \ln D$ are the errors in height and ln D respectviely. The median points of the groups have been found and a line, PQ, fitted to these points.

3. RESULTS

(a) The two-station experiment was performed in December 1955. The height differences and the differences in ln D have been plotted on a scatter diagram and median points found for groups of echoes by a similar method to that described above. The median plots are plotted in Fig. 2 together with a least squares line passing through the origin. The gradient of the line is $8{\cdot}9 \pm 0{\cdot}8$ km. Considerable weight must be given to this determination as it is less dependent on the effect which caused the systematic differences in the slopes determined by Greenhow and Neufeld and by Murray.

(b) The results for the single-station experiment were obtained during fifteen consecutive days of January–February 1959. They have been divided into eight three-hourly groups and plotted on scatter diagrams similar to Fig. 1. The gradients, found by the method described above are presented in column (a) of Table 1.

† *Note added in proof*: A full least-squares solution has since been made using a Mercury digital computor. The gradients presented in this paper have been confirmed by this later analysis.

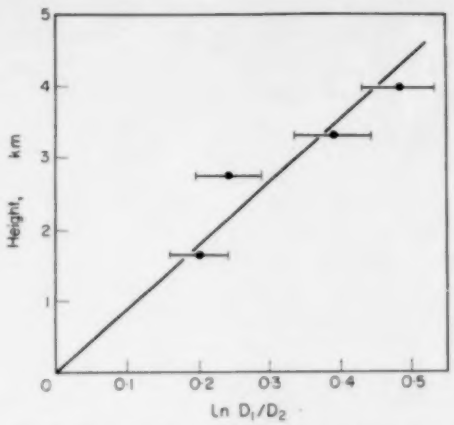


Fig. 2. The variation in the differences in ln D with the height differences for two points on a single meteor trail.

The mean value of the gradients is $(7{\cdot}8 \pm 0{\cdot}3)$ km which is in reasonable agreement with that of the two station experiment given above. A weighted mean value for the two experiments is

Table 1. The Slope of the height ln D relation

Time	Gradient $\frac{d(\text{Height})}{d(\ln D)}$		(km)
(U.T.)	(a)	(b)	(c)
00hr–03hr	8·2	6·6	12·5
03hr–06hr	8·0	6·5	12·5
06hr–09hr	7·3	5·5	12·2
09hr–12hr	8·3	6·7	12·4
12hr–15hr	7·6	5·7	12·6
15hr–18hr	7·2	6·0	14·7
18hr–21hr	7·9	6·6	14·4
21hr–24hr	8·2	6·6	14·6
Means	7·8	6·3	13·2
Greenhow and Neufeld[3]		6·5	
Murray[4]		5·6	30 29 25 19

(a) Assuming errors in both height and ln$D$.
(b) Assuming errors in height only.
(c) Assuming errors in ln$D$ only.

$8{\cdot}0 \pm 0{\cdot}3$ km. The observations refer to a height of 98 km. The significance of a possible diurnal variation in the values of the gradient is discussed in another paper (Greenhow and Hall[12]).

4. DISCUSSION

(a) *Comparison with other results*

Greenhow and Neufeld[3] fitted a line to a height/ln D scatter diagram by dividing the points into groups of similar diffusion coefficient and finding the median height for each group. The line fitting these median points is the best fit to the scatter diagram only when the error in diffusion coefficient is negligible. The line that was obtained gives the most probable height of a meteor of a given decay time, but does not necessarily show the true relation between height and diffusion coefficient. Their value for the slope of the height/ln D relation is approximately 0·8 times the value determined in the present paper.

Murray[4], on the other hand, found the slope for the Adelaide results by a method of least squares on the assumption that there was a negligible height error. However, even if this assumption is justified, the results are misleading since there appears to be a very strong selection effect in the values of ln D measured from the Adelaide observations. The individual values of ln D published by Murray have been compared with the present Jodrell Bank values. Fig. 3(a) is a histogram showing the frequency of occurrence of values of ln D for meteor heights above 100 km taken from the Adelaide results, and Fig. 3(b) is a similar histogram for the Jodrell Bank echoes. It is clear that the Adelaide results are limited to values of ln D below 11·8, i.e. $D < 1{\cdot}35 \times 10^5$ cm^2 sec^{-1}. This limitation is probably due to the method of determining the echo decay time from the Doppler display of the upper atmospheric winds equipment (Weiss[13]). For a decay time to be measurable by this technique, the meteor must last long enough for at least one Doppler cycle to be completed. With wind velocities of 50 m/sec, the maximum value of diffusion coefficient that can be measured is about $1{\cdot}5 \times 10^5$ cm^2 sec^{-1} which is close to the limit actually observed. The selection effect is so marked that the median value of the diffusion coefficient measured at Jodrell Bank lies above the cut-off in the Adelaide distribution. The effect of this selection against meteors with rapid decay is to increase the slope of the height/log$_e$ D relation.

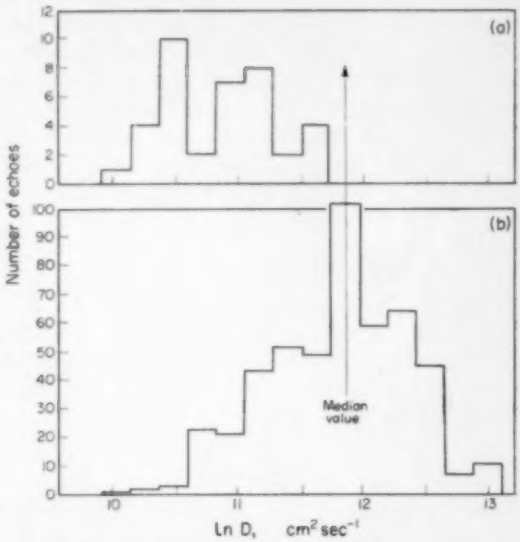


Fig. 3(a). The frequency of occurrence of values of log$_e$D for meteors at heights above 100 km for the Adelaide results (Murray[4]).
(b). The frequency of occurrence of values of log$_e$D for meteors at heights above 100 km for the Jodrell Bank results.

In order to compare the present Jodrell Bank results directly with those discussed above, two additional lines have been fitted to each of the eight three-hourly groups, one on the assumption that all the error occurs in the heights and the other on the assumption that there is only error in the values of the diffusion coefficient. The gradients of these lines are given in Table 1 columns (b) and (c) and may be compared with the mean value found by Greenhow and Neufeld and the values given by Murray. Even when the present observations, which have an estimated error of ± 4 km in height, are analysed in the same way as the Adelaide results, there is still a very large difference in the measured slopes. This discrepancy is due to the selection occurring in the Adelaide observations.

(b) *The effect of inaccuracy in the estimation of errors*

It can be seen from columns (b) and (c) of Table I that the gradient of the height/ln D profile, obtained from measurements on radio meteors, is strongly dependent on the assumptions made about the errors in height and decay time. If it is assumed that height errors are negligible compared with errors in D (column (c)) then the calculated gradient is about twice that which is obtained on the opposite assumption (column (b)). In practice the errors are comparable, and it is necessary to decide their relative importance in order to find the correct value of H_D. The parameter which is used in the calculation is $W=(\Delta h/\Delta \ln D)^2$. The results presented in this paper, Table 1(a), have been found on the assumption that $W=200$, for which $H_D=7{\cdot}8\pm0{\cdot}3$ km..

It is interesting to enquire to what extent the adopted value of W affects the calculated slope of the height/ln D profile. The analysis has therefore been repeated for values of W differing by a factor of 2, when for $W=150$ the gradient $H_D=8{\cdot}9$ km, and for $W=300$ $H_D=7{\cdot}3$ km. It is believed that these values of W represent the limits that could be assumed due to the uncertainty in the errors of height and ln D. Therefore H_D lies between 7·3 and 8·9 km with a most probable value of 7·8 km. The error produced by this effect is greater than that due to the scatter of the individual results and it is most important to include it when an estimate is made of the overall error. Using this value of H_D and also the result of the two station experiment it is concluded that $H_D=8{\cdot}0\pm0{\cdot}7$ km.

(c) *The relation between H_D and the atmospheric scale height H*

The value of 8·0 km for the slope of the diffusion coefficient height relation H_D, is about 20 per cent greater than the generally accepted value of the atmospheric scale height, H (Rocket Panel[14]). An examination of the assumptions involved in the derivation of equation (2) will therefore be made in an attempt to account for the discrepancy.

(i) *Isothermal atmosphere.* The atmosphere in the meteor region is not isothermal and the temperature probably increases with height. Since meteors occur over a limited height range, a sufficient approximation is to assume a linear variation of scale height, H, with height, h. The gradient H_D of the height/ln D line is then given by $H/(1+1{\cdot}5\beta)$ where β is defined by $H=H_0+\beta(h-h_0)$, h_0 being the height at which the scale height is H_0. H_D is therefore reduced by a factor $(1+1{\cdot}5\beta)$ below the average atmospheric scale height in the region. The observations indicate a gradient 20 per cent greater than H and to explain this, it is necessary to postulate a scale height decreasing with height at a rate of about 0·2 km/km. This is highly improbable.

(ii) *The effect of the dissociation of oxygen.* At a height of 95 km molecular oxygen is partially dissociated by ultra-violet light to form atomic oxygen (Mitra[15]). If dissociation is to be responsible for the increase in H_D, then, over a height range from 90 km to 110 km, the effective diffusion coefficient must be reduced by a factor of two relative to the value it would have without dissociation. For 100 per cent dissocation at 110 km, this would require that either the collision cross-section between meteor ions and atomic oxygen is four times that between meteor ions and molecular oxygen, or that the collision cross-section between electrons and atomic oxygen is one hundred times that between electrons and molecular oxygen. As the dissociation is probably less than 100 per cent at 110 km, then the atomic cross-section must be even larger. There is no experimental evidence for either of these increased cross-sections (Klein and Brueckner[16]) and it seems unlikely that the discrepancy in the gradients is due to oxygen dissociations.

(iii) *Latitude or seasonal variations.* Since neither of the above considerations gives a reason why H_D should differ from the atmospheric scale height, H, it may be assumed that $H_D=H$. If this is the case, there is a real difference between the scale height measured during summer over New Mexico at latitude 33°N and that measured in winter at Jodrell Bank (53° 14′ N). This difference may be due to either a latitude or a seasonal variation but

these cannot be distinguished until further analysis of observations made during the summer months is completed. Some support is given to this hypothesis by the results of rocket flights in winter over Fort Churchill at 59° N when, for the height of 130 km, scale heights twice the values given by the Rocket Panel were obtained (Horowitz and Lagow[17]).

Acknowledgements—The research work reported in this paper has been performed at Jodrell Bank Experimental Station under the direction of Professor A. C. B. Lovell, F.R.S. It has been supported in part by the Air Research and Development Command, United States Air Force, contract no. AF61(514)-948.

REFERENCES

1. T. R. Kaiser, *Phil. Mag.* Supp. **2,** 495 (1953).
2. N. Herlofson, *Reports on Progress in Physics* Vol. 11, p. 444 Physical Society London (1948).
3. J. S. Greenhow and E. L. Neufeld, *J. Atmos. Terr. Phys.* **6,** 133 (1955).
4. E. L. Murray, *Planet. Space Sci.* **1,** 125 (1959).
5. J. S. Greenhow and E. L. Neufeld, *Mon. Not. R. Astr. Soc.* **117,** 359 (1957).
6. J. A. Clegg and I. A. Davidson, *Phil. Mag.* **41,** 77 (1950).
7. E. R. Billam and I. C. Browne, *Proc. Phys. Soc.* B **69,** 98 (1956).
8. M. Loewenthal, *Mass Inst. Tech. Lincoln Lab. Tech. Rep.* 132 (1956).
9. T. R. Kaiser and R. L. Closs, *Phil. Mag.* **43,** 1 (1952).
10. J. S. Greenhow, *Proc. Phys. Soc.* B **65,** 169 (1952).
11. J. S. Greenhow and J. E. Hall, *Mon. Not. R. Astr. Soc.* (In Press) (1960).
12. J. S. Greenhow and J. E. Hall, *J. Atmos. Terr. Phys.* (In Press) (1960).
13. A. A. Weiss, *Aust. J. Phys.* **8,** 279 (1955).
14. Rocket Panel, *Phys. Rev.* **88,** 1027 (1952).
15. S. K. Mitra, *The Upper Atmosphere.* The Asiatic Society, Calcutta (1952).
16. M. M. Klein and K. A. Breuckner, *Phys. Rev.* **111,** 1115 (1958).
17. R. Horowitz and H. E. Lagow, *J. Geophys Res.* **63,** 759 (1958).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 115–121. Printed in Great Britain.

SPUTTERING AS A POSSIBLE MECHANISM FOR INCREASE OF IONIZATION IN THE VICINITY OF LOW-ALTITUDE SATELLITES*

G. D. MAGNUSON and D. B. MEDVED

Physics Section of Convair, A Division of General Dynamics Corporation, San Diego, California

(*Received* 19 *May* 1960)

Abstract—It is suggested that for altitudes below 200 km there may be sufficient sputtering produced by ion, atomic, and molecular impact on the vehicle surface to lead to some increase in ionization above ambient at distances of one mean path or less from the vehicle. The sputtering process injects into the enviroment atoms of metallic elements whose ionization potentials are 1/2 to 1/3 those of the ambient species. Collision energies in the centre of mass system vary from ionization thresholds (E_i) to 2 to 3 times E_i. The resulting ionization in the surroundings may then be calculated if values for β, the probability of any collision leading to ionization, is known. On the basis of the crude model presented, β should be greater than or equal to 10^{-6} for this mechanism to be of interest.

1. INTRODUCTION

Early in October 1957, J. D. Kraus[1] of Ohio State University made some interesting observations on what he believed to have been the last days in orbit of Sputnik I by using a simple CW reflection technique. According to his description the 20 Mc signal was very rough and unsteady during the satellite's approach. However, sometime before the satellite reached its closest point of approach the signal suddenly became very smooth and steady and remained so during the rest of the pass. This effect was observed on a number of passes with the transition from rough-to-smooth signal occurring when the angle between the satellite path and a line from the satellite to Columbus, Ohio, was about 60 degrees. The most interesting facet of the observations was the fact that the effect was observed only while the satellite was approaching and not at all while it was receding.

It was suggested by J. Kraus that the observations were due to an ionization column moving ahead of the satellite, and originated in the following manner. In a perfectly elastic collision between a satellite moving with velocity V and a stationary gas particle, the particle will be reflected from the satellite with a relative velocity V and, hence, with a velocity $2V$ relative to the stationary ambient gas particles. Kraus then suggests that these reflected particles are moving fast enough to suffer ionization in collisions with the stationary particles, resulting in a column of ionization moving ahead of the satellite. He also suggests the possibility of reflecting ionized particles back into ambient thus doubling the ionization concentration. However, even if the sticking coefficient of impacting electrons on the surface were zero for a spherical surface, the ionization column ahead of the satellite would have an increase above ambient of considerably less than 100 per cent.

It is our purpose to suggest sputtering of the satellite's surface as an alternative mechanism for this increase of ionization in the vicinity of a low-altitude satellite. For a satellite moving with velocity V the energy with which a stationary gas particle impacts the satellite's surface is:

$$E_{im} = 1/2 mV^2 \text{ (erg)} \qquad (1)$$

where m is the mass of the gas particle in grams and V is the velocity in cm/sec. To calculate

* This work was supported by ARPA Order No. 4259, Contract AFCRC-AF-19(604)5554.

Table 1. Energetics of the Collision Processes
(Energies are in electron volts)

	$V_1 = 7{\cdot}3 \times 10^5$ cm/sec					$V_2 = 7{\cdot}9 \times 10^5$ cm/sec				
	M	M_r	E_{subl}	E_{sc}	E_{ion}	M	M_r	E_{subl}	E_{sc}	E_{ion}
N_2	28	14			15·58	28	14			15·58
Cu	63·5	19·43	3·56	11·89	7·70	63·5	19·43	3·56	13·08	7·70
Ti	48	17·68	4·62	10·78	6·5	48	17·68	4·62	14·83	6·5
C	12	8·40	7·45	3·90+	11·22	12	8·40	7·45	7·37+	11·22
Sn	119	22·667	3·38	11·66	7·3	119	22·667	3·38	14·06	7·3
O	16	8·00			13·61	16	8·00			13·61
Cu	63·5	12·78	3·56	5·28+	7·70	63·5	12·78	3·56	6·79+	7·70
Ti	48	12	4·62*		6·5	48	12	4·62	5·50+	6·5
C	12	6·86	7·45*		11·22	12	6·86	7·45*		11·22
Sn	119	14·10	3·38	5·36+	7·3	119	14·10	3·38	6·75+	7·3
A	40	20			15·76	40	20			15·76
Ti	48	21·8	4·62	17·37	6·5	48	21·8	4·62	21·21	6·5
Cu	63·5	24·5	3·56	18·56	7·70	63·5	24·5	3·56	22·33	7·70

	E_{im}	E_{ke}		E_{im}	E_{ke}
N_2	7·74	15·48+	N_2	9·07	18·14
O	4·43	8·86+	O	5·18	10·36+
A	11·06	22·12	A	12·96	25·92

	$V_3 = 1{\cdot}1 \times 10^6$ cm/sec				
	M	M_r	E_{subl}	E_{sc}	E_{ion}
N_2	28	14			15·58
Cu	63·5	19·43	3·56	30·97	7·70
Ti	48	17·68	4·62	30·44	6·5
C	12	8·40	7·45	24·60	11·22
Sn	119	22·667	3·38	29·35	7·3
O	16	8			13·61
Cu	63·5	12·78	3·56	15·81	7·70
Ti	48	12	4·62	15·29	6·5
C	12	6·86	7·45	10·86+	11·22
Sn	119	14·10	3·38	14·94	3·3
A	40	20			15·76
Ti	48	21·8	4·62	45·63	6·5
Cu	63·5	24·5	3·56	46·41	7·70

	E_{im}	E_{ke}
N_2	17·58	35·16
O	10·05	20·10
A	25·12	50·24

E_{subl} = Sublimation energy of the solid.

E_{ion} = Ionization energy of the atom or molecule.

E_{sc} = Collision energy in the center of mass system of a sputtered atom and a stationary atmospheric constituent.

E_{ke} = Collision energy in the center of mass system of an elastically reflected gas atom and an identical stationary atmosphere atom.

E_{im} = Energy of impact of the atmospheric constituent on a vehicle.

M = Atomic mass.

M_r = Reduced atomic mass.

$V_{1,2,3}$ = Missile velocity.

*No physical sputtering will occur

+Below ionization threshold

the energy in eV,

$$E_{im} = 5{\cdot}19 \times 10^{-13} MV^2 \qquad (2)$$

where M is the mass of the particle in atomic mass units. For the usual atmospheric constituents E_{im} is in the range from about 4 to 10 eV. Until recently this energy range would have been considered too low to sputter a satellite surface since threshold energy measurements indicated threshold energies of 30 eV or higher. Recent Russian work[2], however, has shown that threshold energies are in the range of 10 eV or less. In fact, since the Russian work was done at perpendicular incidence and the sputtering threshold energy should decrease as the ion's angle of incidence increases it is quite likely that sputtering of a satellite surface may occur at impacting energies as low as the sublimation energy of the solid. These sputtered atoms are moving with at least the velocity V of the missile with respect to the ambient gas particles and, in many cases, have kinetic energies in the reduced mass system higher than ionization threshold. Sputtering is thus offered as an alternative mechanism for a possible ionization increase. The key factor to be kept in mind throughout the following discussion is that sputtering injects atoms of metallic elements whose ionization potentials are roughly 1/2 to 1/3 those of the ambient species such as A, O_2, etc.

Table 1 contains a summary of some calculations made for two different vehicle velocities and two different ranges of altitude. The two velocities $7{\cdot}3 \times 10^5$, and $7{\cdot}9 \times 10^5$ cm/sec correspond to ICBM and satellite velocities, respectively. Calculations are given for the impacting species O, N_2 and A. Nitrogen molecules are the dominant species, with regard to relative concentrations, in the altitude range 120 to 200 km and oxygen atoms are the dominant species above 260 km altitude. Argon is included because of the high impact energies involved. The term E_{ke} is the collision energy in the centre of mass system for an ambient particle elastically scattered from the satellite colliding with a similar, stationary gas particle (where we assume that the surface is a flat plate and hence the relative velocity is 2V). Then

$$E_{ke} = 1/2 m_r (2V)^2 = 2 m_r V^2 \text{ erg} \qquad (3)$$

where m_r is the reduced mass in grams of the two colliding particles. Since, for the Kraus collision model, $m_r = m/2$,

$$E_{ke} = 2\frac{m}{2} V^2 = mV^2 = 2E_{im}. \qquad (4)$$

The collision energy in the centre of mass system between a sputtered atom and a stationary gas particle is E_{sc}.

$$E_{sc} = 1/2 m_r V^{*2} \qquad (5)$$

where

$V^* = V + V_s$ = velocity of sputtered atom relative to the ambient. (6)

$V_s = \sqrt{\dfrac{2(E_{im} - E_{subl})}{m_s}}$ = the velocity of sputtered atom relative to the satellite, and

m_s = mass of the sputtered atom.

Table 1 thus exhibits the data for three atmospheric species N_2, O and A, in collision with atoms of Cu, Ti, Sn and C.

In Table 2, we have summarized the densities of atmospheric species, mean free paths λ and electron densities. For the altitude range 80 to 220 km N_2 is the dominant species whereas above 240 km O is the most prevalent. We shall use Tables 1 and 2 for comparison of the Kraus reflection–collision model and the sputtering mechanism for increase of ionization. In calculating E_{ke} and E_{sc} we have considered the best possible case; it is assumed the velocity of the reflected or sputtered atoms adds algebraically to the satellite velocity. This is equivalent to assuming a stationary atmosphere ($T = 0°$K) with a flat plate moving through it with the reflected or sputtered particles being emitted in a specular manner. For an observer in the reference frame of the ambient the situation is summarized in Fig. 1.

2. CALCULATION OF IONIZATION PRODUCED BY NEUTRAL-NEUTRAL IMPACT

Flat plate case

Consider a flat plate having an area A and moving with velocity V. Then the number of

Table 2. Summary of Upper Atmosphere Parameters

Alt (km)	λ cm	Densities of species, ρ (No./cm³)					Electron Density n_e (No./cm³)		$\alpha_n = \frac{n_e}{\rho}$
		N_2	O_3	N	O	A	Day	Night	
80	$4{\cdot}18\times10^{-1}$	$3{\cdot}52\times10^{14}$	$9{\cdot}43\times10^{13}$	$2{\cdot}09\times10^{10}$	$2{\cdot}15\times10^{11}$	$4{\cdot}17\times10^{12}$	$8{\cdot}0\times10^{2}$	$1{\cdot}0\times10^{9}$	$1{\cdot}4\times10^{-8}$
100	$1{\cdot}14\times10^{1}$	$1{\cdot}16\times10^{13}$	$1{\cdot}40\times10^{12}$	$6{\cdot}02\times10^{9}$	$3{\cdot}40\times10^{12}$	$1{\cdot}37\times10^{11}$	$4{\cdot}5\times10^{4}$	$1{\cdot}5\times10^{4}$	$8{\cdot}0\times10^{-7}$
120	$1{\cdot}79\times10^{2}$	$6{\cdot}95\times10^{11}$	$3{\cdot}0\times10^{10}$	$2{\cdot}28\times10^{9}$	$3{\cdot}13\times10^{11}$	$8{\cdot}23\times10^{9}$	$1{\cdot}22\times10^{5}$	$8{\cdot}4\times10^{4}$	$6{\cdot}0\times10^{-6}$
140	$1{\cdot}77\times10^{3}$	$6{\cdot}94\times10^{10}$	$2{\cdot}62\times10^{9}$	$9{\cdot}75\times10^{8}$	$6{\cdot}19\times10^{9}$	$1{\cdot}63\times10^{8}$	$9{\cdot}5\times10^{4}$	$1{\cdot}2\times10^{5}$	$2{\cdot}5\times10^{-5}$
160	$8{\cdot}78\times10^{3}$	$1{\cdot}38\times10^{10}$	$3{\cdot}41\times10^{8}$	$4{\cdot}24\times10^{8}$	$1{\cdot}90\times10^{9}$	$5{\cdot}01\times10^{7}$	$1{\cdot}20\times10^{5}$	$1{\cdot}65\times10^{5}$	$8{\cdot}0\times10^{-5}$
180	$2{\cdot}81\times10^{4}$	$4{\cdot}23\times10^{9}$	$7{\cdot}73\times10^{7}$	$2{\cdot}41\times10^{8}$	$1{\cdot}01\times10^{9}$	$1{\cdot}27\times10^{7}$	$2{\cdot}40\times10^{5}$	$2{\cdot}4\times10^{5}$	$2{\cdot}7\times10^{-4}$
200	$6{\cdot}60\times10^{4}$	$1{\cdot}55\times10^{9}$	$2{\cdot}51\times10^{7}$	$1{\cdot}45\times10^{8}$	$5{\cdot}82\times10^{8}$	$3{\cdot}78\times10^{6}$	$3{\cdot}7\times10^{5}$	$3{\cdot}7\times10^{5}$	$8{\cdot}0\times10^{-4}$
220	$1{\cdot}36\times10^{5}$	$6{\cdot}41\times10^{8}$	$9{\cdot}27\times10^{6}$	$9{\cdot}28\times10^{7}$	$3{\cdot}54\times10^{8}$	$1{\cdot}28\times10^{6}$	$5{\cdot}8\times10^{5}$	$5{\cdot}8\times10^{5}$	$2{\cdot}5\times10^{-3}$
240	$2{\cdot}51\times10^{5}$	$2{\cdot}90\times10^{8}$	$3{\cdot}81\times10^{6}$	$6{\cdot}18\times10^{7}$	$2{\cdot}25\times10^{8}$	$4{\cdot}78\times10^{5}$	$1{\cdot}1\times10^{6}$	$1{\cdot}1\times10^{6}$	$8{\cdot}3\times10^{-3}$
260	$4{\cdot}36\times10^{5}$	$1{\cdot}42\times10^{8}$	$1{\cdot}70\times10^{6}$	$4{\cdot}26\times10^{7}$	$1{\cdot}49\times10^{8}$	$1{\cdot}95\times10^{5}$	$2{\cdot}2\times10^{6}$	$2{\cdot}2\times10^{6}$	$2{\cdot}3\times10^{-2}$
280	$7{\cdot}04\times10^{5}$	$7{\cdot}36\times10^{7}$	$8{\cdot}16\times10^{5}$	$1{\cdot}90\times10^{7}$	$1{\cdot}02\times10^{8}$	$8{\cdot}58\times10^{4}$	$4{\cdot}0\times10^{6}$	$4{\cdot}0\times10^{6}$	
300	$1{\cdot}08\times10^{6}$	$4{\cdot}03\times10^{7}$	$4{\cdot}15\times10^{5}$						

atoms sputtered per cm²/sec is

$$S=\mu J=\mu\rho V \tag{7}$$

where J is the incident particle flux density, ρ is the ambient particle volume density and μ is the sputtering yield. The total number of atoms sputtered in a time dt is then

$$SAdt=\mu\rho VAdt \tag{8}$$

However, the plate sweeps out a volume $AVdt$ in this same time and hence the volume density of sputtered atoms (in distances less than a mean free path from the plate) is

$$\rho_s=\mu\rho. \tag{9}$$

In a distance, one mean free path, λ (assumed to be the same for both sputtered and ambient particles) from the plate $(1-e^{-1})$ of the sputtered atoms will collide with ambient. If a fraction, β, of these are assumed to be ionizing collisions then within a distance λ of the vehicle, a density of ionization caused by the sputtered atoms will be

$$\Delta n_e=\mu\rho\left(1-\frac{1}{e}\right)\beta \cong \mu\beta\rho \tag{10}$$

The total electron density is given by

$$n_e=n_0+\Delta n_e =(\alpha+\mu\beta)\,\rho \tag{11}$$

where α represents the degree of ionization of the ambient atmosphere. In order for sputtering to produce some effect on electron densities, it is then necessary that

$$\mu\beta \gtrsim \alpha \tag{12}$$

A plot of α as a function of altitude is given in Fig. 2 using the data of Table 2. The values for α are those for percentage of ionization corresponding to night time, $\alpha=\alpha_n$. On the same plot the intercepts of the α_n curve with various values of the product $\mu\beta$ are shown. Since μ is expected to be no larger than 10^{-2} (and more probably is 10^{-3} or 10^{-4}) we see that the altitude domain of interest probably corresponds to a range of altitude 100 to 200 km.

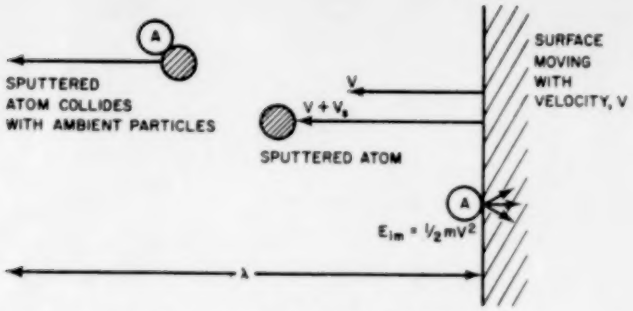


Fig. 1. Sputtering and collision processes.

Spherical vehicle, the effect of the Knudsen number

In the previous derivation of the equation $\Delta n_e = \mu\beta\rho$ it was assumed that the ionization produced in a mean free path essentially was equivalent to an equilibrium concentration. However, depending on the distribution function of the ejected electrons and their velocities relative to vehicle velocity, the actual Δn_e may be considerably smaller. In the following discussion, an attempt is made to modify the equation by taking into account an isotropic sputtering source in the coordinate system of an observer on the vehicle (Fig. 3) where the beam of sputtered atoms and free electrons resulting from neutral-neutral collisions is not confined to the volume swept out by the object. In addition, we now assume that the sputtered particles have some small but finite velocity relative to the surface from which they were emitted. The incident number of impacting particles per second is given by $2\pi R^2\rho V$ and in a time dt, the emitted number of sputtered atoms is

$$N_s = \mu\rho V 2\pi R^2 dt \tag{13}$$

The volume concentration of sputtered atoms is thus,

$$\rho_s]_{n.c.} = \frac{N_s}{\text{Vol}} = \frac{\mu\rho V\,(2\pi R^2)\,dt}{4\pi x^2 V_s\,dt^1}$$

where $4\pi x^2 V_s dt$ is the volume of a spherical shell (Fig. 3) swept out by sputtered atoms with velocity V_s* in a time interval dt^1 equivalent to dt. Thus,

$$\rho_s(x)]_{n.c.} = \frac{\mu\rho}{2}\left(\frac{R}{x}\right)^2 \frac{V}{V_s} \tag{14}$$

Equation (14) which gives the concentration of sputtered atoms is derived without considering collisions. A considerable simplification may be introduced in the following manner. Assume that for the energies of collision in the reduced mass system β is such a strongly decreasing

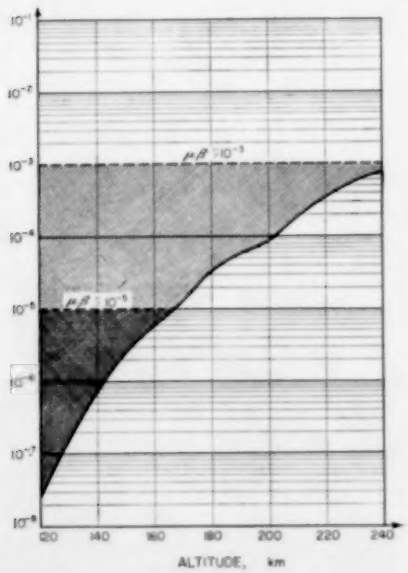


Fig. 2. Fraction of ionization α_n as a function of altitude compared to $\mu\,\beta$ product.

* In this preliminary treatment V_s is considered to be independent of x, i.e., within distances of several mean free paths the velocity is considered constant.

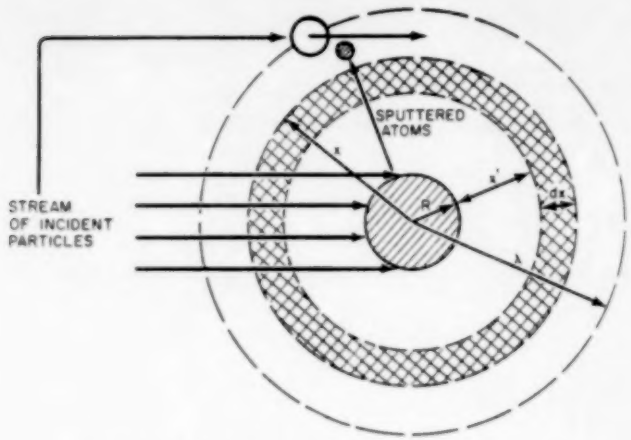


Fig. 3. Spherical distribution of sputtered atoms.

function of energy that only in the first collision is it non-zero; i.e. ionization is produced only during the first collision. Equation (14) must now be modified to take into account collisions in the usual manner, i.e.

$$\rho_s(x) = \frac{\mu\rho}{2}\left(\frac{R}{x}\right)^2 \frac{V}{V_s} \exp -(x-R)/\lambda \tag{15}$$

where λ is an "effective" mean free path. The mean free path is anisotropic due to the directed flow of ambient along the x-axis. Intuitively, λ is a minimum for V and V_s anti-parallel. It can be shown that:

$$\lambda = \frac{V_s}{[V_s^2 + V^2 - 2VV_s\cos\phi]^{1/2}}\lambda_0 \tag{16}$$

Thus, the electron concentration in a shell dx thick is given by

$$dn_e = \frac{\beta\mu\rho}{2}\left(\frac{R}{x}\right)^2 \frac{V}{V_s} \exp -(x-R)/\lambda \frac{dx}{\lambda}$$

This result is obtained by considering that ρ_s atoms reach the shell and that of these dx/λ collide in the distance x to $x+dx$. Of these collisions, a fraction β are ionizing ones.

From the standpoint of a stationary problem we can obtain a more explicit answer in the following manner. Consider that the rate of formation of electrons is given by:

$$\frac{dn_e}{dt} = \beta\rho_s V_s/\lambda$$

and that the recombination rate is characterized by a mean lifetime τ. Then at equilibrium

$$n_e = \frac{dn_e}{dt}\tau = \beta\rho_s(V_s/\lambda)\tau(n_e)$$

and finally

$$n_e(x) = \beta\mu\rho\left(\frac{R}{x}\right)^2 \frac{V\tau}{\lambda} \exp -(x-R)/\lambda \tag{17}$$

We may calculate a typical case as follows. Consider $\tau^{-1} = n_e V_e \sigma_e$ where σ_e is the recombination cross-section and V_e is the electron velocity here assumed thermal and approximately 10^7 cm/sec. At 140 km, n_0, the ambient electron density is approximately 10^5 cm^{-3} and $\lambda_0 \sim 10^5$ cm, with $\rho \sim 10^{11}$. Assuming $\sigma_e \sim 10^{-12}$ cm^2 the mean lifetime of the electrons generated by the collisions is $\tau = 10^5/n_e$.

From inspection of equation (17) the value of x which corresponds to equality of n_0 and n_e is $x \sim 10^3$ cm; i.e. for distances of one mean free path from a 1-meter-diameter sphere the electron density is doubled at 140 km altitude.

This result, however, has been obtained by using a lifetime which corresponds to an accumulation time of electrons at a distance x from the vehicle. The source term is changing

as a function of time since the vehicle moves many free paths in the mean lifetime calculated. Thus, the time used for τ in equation (17) is incorrect and should be replaced by something on the order of the time it takes the vehicle to move one diameter. On the other hand, it is evident that neglecting diffusion of the electrons and ions produced, a long train or wake is established since the electrons produced at a distance x from the vehicle remain there in the absolute reference frame as the vehicle moves on. Thus, the length of the trail can be considered as the distance covered by the vehicle during a time interval corresponding to the mean lifetime of the electrons. For $\tau = 1$ sec, a train of high density (corresponding to electron concentrations for $x \sim R$) extending as much as 10 km behind the object could appear. However, the diffusion mechanism would probably decrease this distance considerably. In essence we have recovered our original equation (10) as valid for large distances, choosing $x = R$ and $V\tau = \lambda$.

3. CONCLUSIONS

It appears that for the altitude region 80 to 120 km which corresponds to the region of so-called "early-shock" formation, that sputtering* could conceivably be a profound and controlling mechanism in the build-up of ionization from ambient densities to those existing in the fully developed shock.

A more complete analysis of the free molecule flow region will depend on the ability to solve the problem of the distribution of sputtered material and electrons and then to calculate the lifetime for readjustment to equilibrium ambient densities of ionization. A first step in this direction will be to set up and solve the diffusion problem of an isotropic point source moving with a high velocity in a medium characterized by a given mean free path. Unfortunately this solution would only be valid for values of x corresponding to many mean free paths. It is the unique feature of the theory considered here that it is essentially a single-collision mechanism, i.e. either electrons are produced during the first collisions or they are not produced at all. The result of the first collision essentially thermalizes the atoms or the resulting ions and electrons and hence they are removed from the problem; i.e. these are no longer properly treated in the reference frame of the moving vehicle and must be considered as remaining in its wake from the standpoint of a stationary observer.

Acknowledgements—The authors wish to express their appreciation for many stimulating and informative discussions with their colleagues in the Physics Section, particularly L. Kraus* and S. Rand, and with W. Teutsch (General Atomic) and F. Hurlbut (Physics Section Consultant from the University of California–Berkeley).

* By introduction of "catalytic" species whose ionization potentials are considerably below those of ambient species into the interaction region.

REFERENCES

1. J. D. Kraus, *Proc. Inst. Radio Engrs.* N.Y. **46**, 1580 (1958).
2. N. D. Morgulis and V. D. Tishchenko, *J. Exper. Theoret. Phys. U.S.S.R.* **30**, 54 (1956).
3. R. M. Rhodes, *Atmosphere Profile* Internal Memorandum Convair, Physics Section Ph-020-M, August 12, 1958.
4. L. E. Miller, *J. Geophys. Res.* **62**, (3)351 (1957).

* Now with Republic Aviation Corp.

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 122–136. Printed in Great Britain.

DETERMINATION OF THE VERTICAL DISTRIBUTION OF OZONE FROM THE MEASUREMENT OF DIFFUSELY REFLECTED ULTRA-VIOLET SOLAR-RADIATION*

Z. SEKERA and J. V. DAVE†
Department of Meteorology, University of California, Los Angeles

(*Received* 21 *July* 1960)

Abstract—The problem of diffuse reflection of solar ultra-violet radiation in the presence of ozone has been solved for a plane-parallel atmosphere of finite optical thickness. The atmosphere is divided into two layers. The scattering optical thickness of the upper layer is small enough to permit one to consider only primary scattering. Ozone is distributed in this layer according to a definite analytical function. The scattering optical thickness of the lower layer is large and all the orders of scattering have been taken into consideration for it. It is assumed that the lower layer does not contain any ozone.

The radiation received by an observer on the top of the atmosphere would consist of two parts. The first part originates from the upper layer as a consequence of scattering of direct solar radiation falling on it. The other part originating from the lower layer consists of four different components. The results of computations presented here contain intensities of each of these five components for several values of ozone absorption and for several directions of incident and scattered radiations. It is expected that this will be helpful in obtaining a broad, semi-quantitative picture of the amount of radiation emerging from each of the layers under different conditions which one is likely to encounter. The results are expected to be of some assistance in further studies related to the problem of the measuring of ozone from satellites or high level balloons. This aspect has also been considered in some detail.

1. INTRODUCTION

The idea of determining ozone from a satellite was put forth by Singer and Wentworth[1] in 1957. Considering primary scattering only, they computed the rate at which the contribution to diffusely reflected radiation along nadir changes as a function of optical depth in the atmosphere. They carried out computations for two different wave lengths (2800 and 3000 Å) and for three different zenith distances of the sun (0°, 30° and 60°). They determined the "effective depth" to which the detector would see the atmosphere, by neglecting the depth below which the rate has a value less than one-tenth of the maximum value at the top. Distributing 0·43 cm of the total ozone according to the photochemical theory so that 0·30 cm of it lies above 15 km, they came to the conclusion that a detector which is sensitive at 3000 Å would effectively see the atmosphere at 24 km when the sun is at zenith. The height increases to 25 and 27 km as the solar distance increases to 30° and 60° respectively. The conditions selected are rather extreme and there are several places on the globe where total ozone amount above 15 km is less than 0.30 cm.

The procedure adopted by Singer and Wentworth would be valid as long as the diffused radiation originates from the part of the atmosphere well above the ozone maximum. However, as the effective scattering depth comes closer to the ozone maximum, a significant portion of solar radiation would reach the part of the atmosphere below the ozone maximum. In this case, one has to take into consideration the contribution from the lower atmosphere to diffusely reflected radiation. Even though the radiation falling on the top of the lower atmosphere is fairly attenuated, because of its large scattering air-mass and lower ozone concentrations, its contribution to the upward radiation can be comparable to that originating from the

* Contributions to Meteorology, Dept. of Meteorology, U.C.L.A., No. 60. The research reported in this paper was supported by the U.S. Air Force, Cambridge Research Center, Contract AF 19(604)-2429).

† Now at Dept. of Meteorology, Imperial College of Science and Technology, London, England.

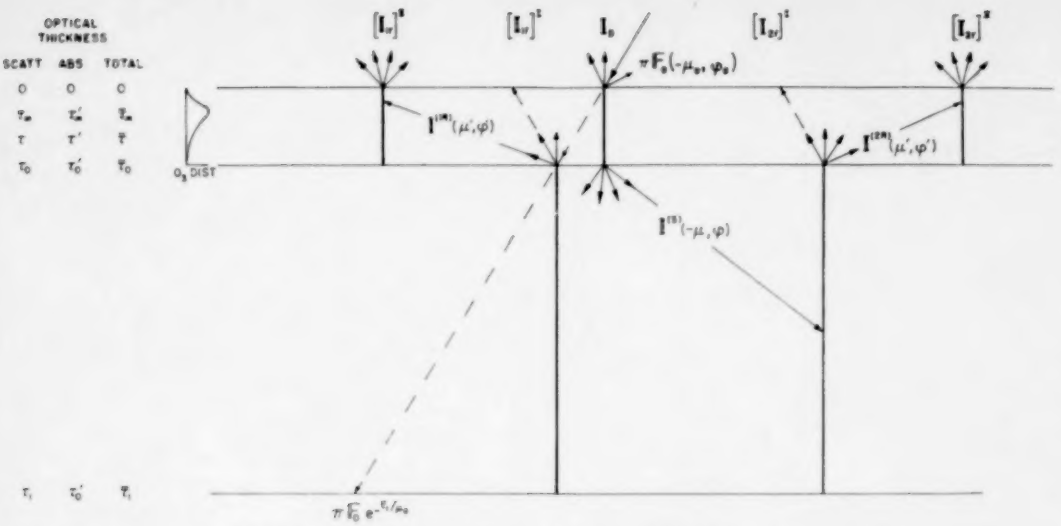


Fig. 1. Different components of upward radiation.

upper layer. As will be seen later, the lower atmosphere is being illuminated from above by reduced, direct solar flux and also by solar radiation diffusely transmitted by the upper layer. The first one depends upon air-mass, total ozone and the zenith angle of the sun, while the latter is determined by these factors as well as by the ozone distribution and by the direction of transmission. Strictly from the physical considerations, one can see that the contribution to upward radiation due to the latter becomes more important as the attenuation of direct solar radiation in the upper layer increases.

The ultra-violet radiation diffusely reflected by the earth's atmosphere is affected to a great extent by surface reflectivity as well as by haze and clouds in the troposphere. Hence, in order to obtain a reliable and instantaneous picture of ozone distribution on a world-wide scale, one must develop a method which is completely insensitive to tropospheric changes, but very sensitive to small changes in the concentration of ozone in the stratosphere. The results can then be used to study different aspects of photochemistry and circulation in the stratosphere.

In the present theoretical study we consider the atmosphere as being plane-parallel, homogeneous and of infinite extent in the horizontal direction but with a finite scattering optical thickness along the vertical. The entire atmosphere is divided into two layers. In the upper layer ozone is assumed to have a distribution as predicted by the photochemical theory. On the other hand, it is assumed that there is no ozone in the lower layer.

The problem has been formulated in such a manner that the contributions from the lower layer can easily be evaluated. The computations have been so performed as to obtain relative magnitudes of the individual components contributing to the upward radiation under different ozone concentrations as well as for different directions of incident and observed radiations. In the case of the component of upward radiation originating strictly from the upper layer, details have been obtained so as to get an idea about the part of the upper layer responsible for scattering under different situations.

Theoretical and computational details have been kept to a minimum here but have been discussed fully elsewhere[2].

2. THEORY

The problem of diffuse reflection and transmission of solar radiation by a plane-parallel,

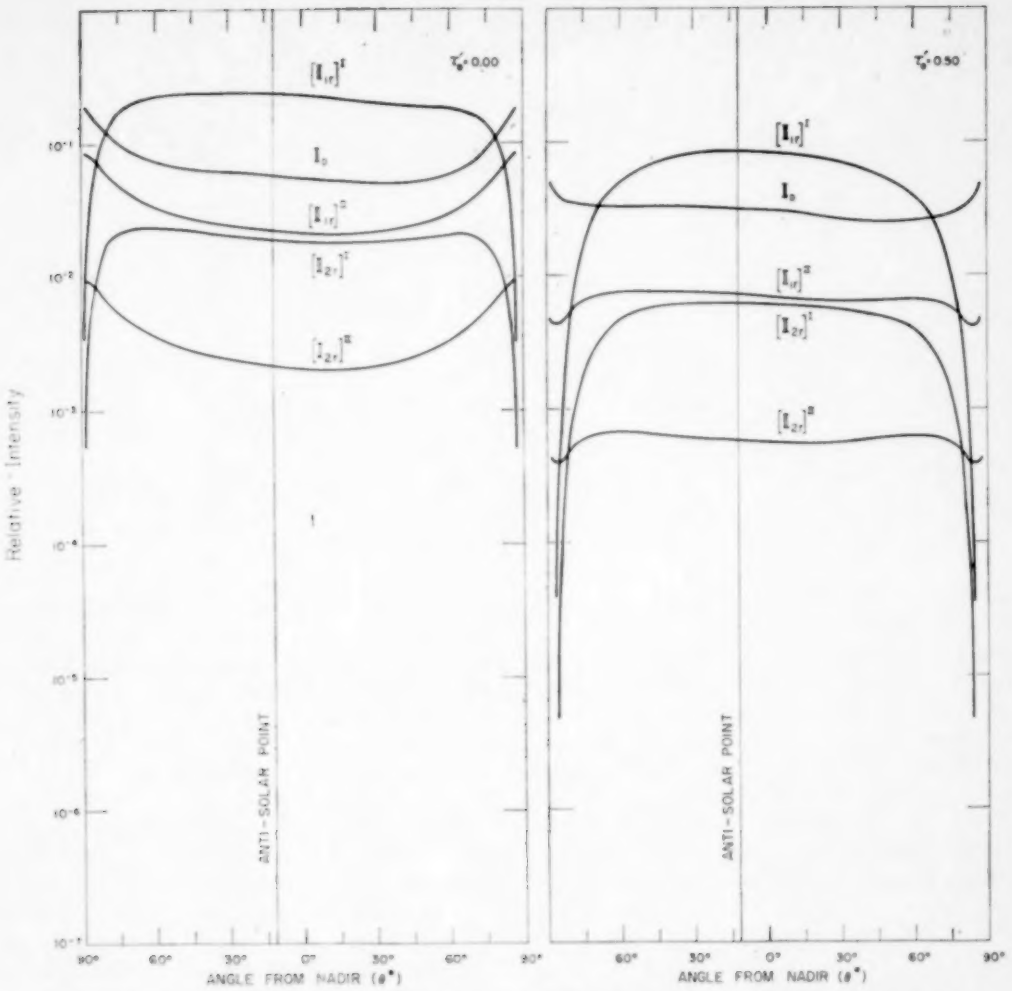


Fig. 2. Relative intensities of five different components of upward radiation for $\mu_0 = 0{\cdot}98$ ($\theta_0 = 11{\cdot}5°$).

homogeneous, infinite atmosphere of finite optical thickness and scattering according to Rayleigh's law has been solved by Chandrasekhar[3]. However, his treatment of the problem and subsequent computations can apply only to the case of perfect scattering in a homogeneous atmosphere. On the other hand, in the problem of diffuse reflection (or transmission) of solar ultra-violet radiation by the earth's atmosphere, one has to take into consideration the absorption due to ozone in the Hartley band. The problem is further complicated due to the fact that the ozone/air ratio is not independent of height. The only favorable factor is that emission by ozone in this region of the spectrum is completely unimportant because of the very low temperatures of the atmosphere.

In order to solve the equation of radiative transfer, so as to take into consideration all the orders of scattering in a case where the absorption and scattering have to be considered, it is necessary to start from the very beginning. This

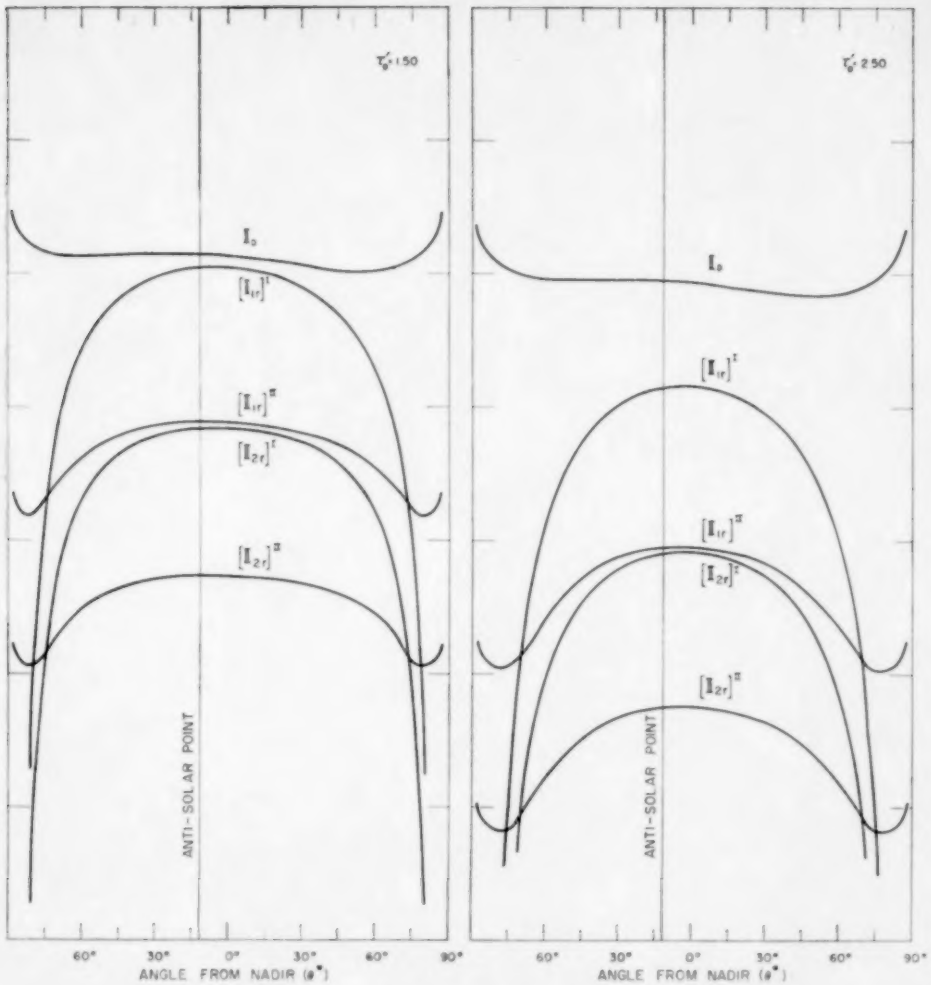


Fig. 2 cont'd.

would demand writing down eight equations based on the "Principle of Invariance" and the formation of some integrals which can be used to obtain a complete solution of the equation of radiative transfer. To the best of our knowledge, no one seems to have embarked on this fundamental problem.

However, for the particular case in which we are interested, we wish to evaluate the contribution to the upward radiation due to the lower layer separately, and this is mainly for the purpose of determining the conditions under which it becomes negligible. In such a case, the lower atmospheric layer can be taken to be a perfect scatterer without introducing any serious error. With this approximation, the upper layer containing ozone has relatively small scattering optical thickness and hence one can consider only primary scattering in this layer. The amount of error involved due to the omission of the second and higher orders of scattering in the upper layer will be discussed later.

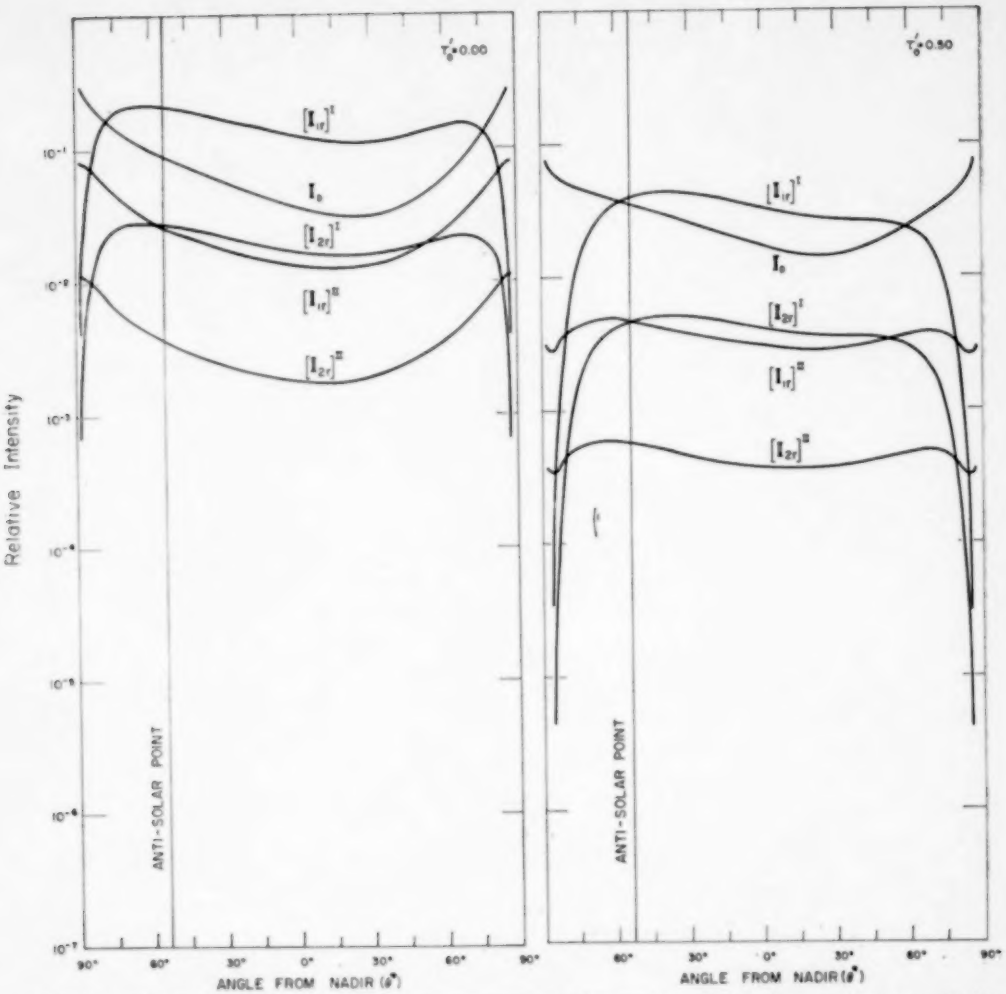


Fig. 3. Relative intensities of five different components of upward radiation for $\mu_0 = 0{\cdot}60$ ($\theta_0 = 53{\cdot}1°$).

Fig. 1 shows the different components of radiation diffusely reflected by a plane-parallel atmosphere divided into two layers. The normal scattering optical thickness of the entire atmosphere is taken to be τ_1. In the upper layer of the normal scattering optical thickness τ_0, ozone is assumed to have been distributed according to the relationship

$$x = \left\{\frac{\kappa}{e}\right\}^{\kappa} \frac{x_0}{H} [\xi \exp(1-\xi)]^{\kappa} \qquad (1)$$

where

x = amount of ozone in cm at N.T.P. in one kilometer layer

κ = a numerical constant between 1·5 and 2·0

e = base of natural logarithms

x_0 = total ozone amount in cm at N.T.P.

H = equivalent scale height

$\xi = \tau/\tau_m = p/p_m$ (if the scattering mass coefficient is assumed constant)

τ_m = normal scattering optical thickness at the point of ozone maximum.

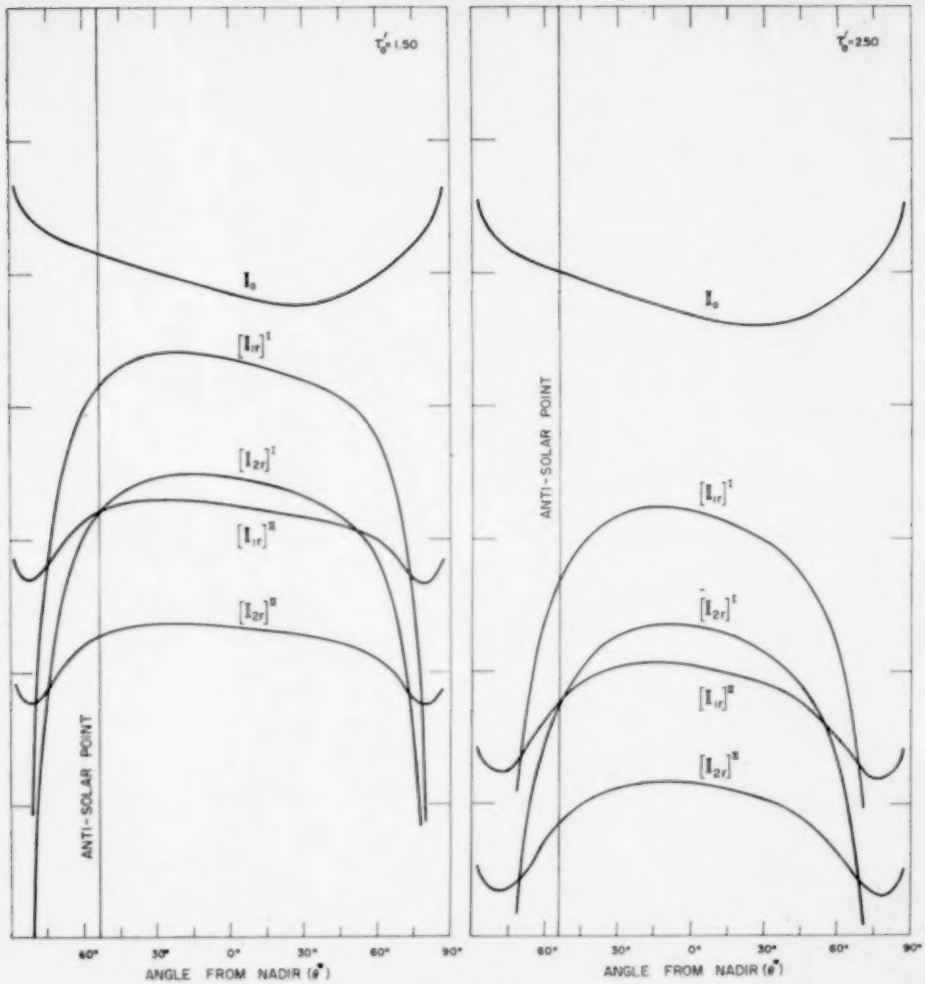


Fig. 3 cont'd.

From this, one can arrive at the formula, giving normal absorption optical thickness (τ') due to ozone as a function of normal scattering optical thickness τ. The total normal optical thickness ($\bar{\tau}$) at any point in the upper layer is given by $\tau+\tau'$. The scattering normal optical thickness of the lower layer is $\tau_1-\tau_0$ and since it is assumed that there is no ozone in the lower layer, total optical thickness ($\bar{\tau}_1$) of the entire atmosphere is $\tau_1+\tau'_0$. As far as possible the notations introduced by Chandrasekar in *Radiative Transfer* are used and hence the radiation is defined by three Stokes parameters, considered as components of an intensity matrix.

Let a parallel beam of radiation of net flux $\pi\mathbf{F}_0$ per unit area normal to itself, be incident on the top of the atmosphere from the direction $(-\mu_0, \varphi_0)$. $\mu_0=|\cos\theta_0|$ where θ_0 is the zenith angle of the sun and φ_0 is the angle which the vertical plane through the sun makes with a vertical plane of reference. The radiation diffusely reflected in the direction $(\mu^*=\cos\theta^*, \varphi^*)$, by

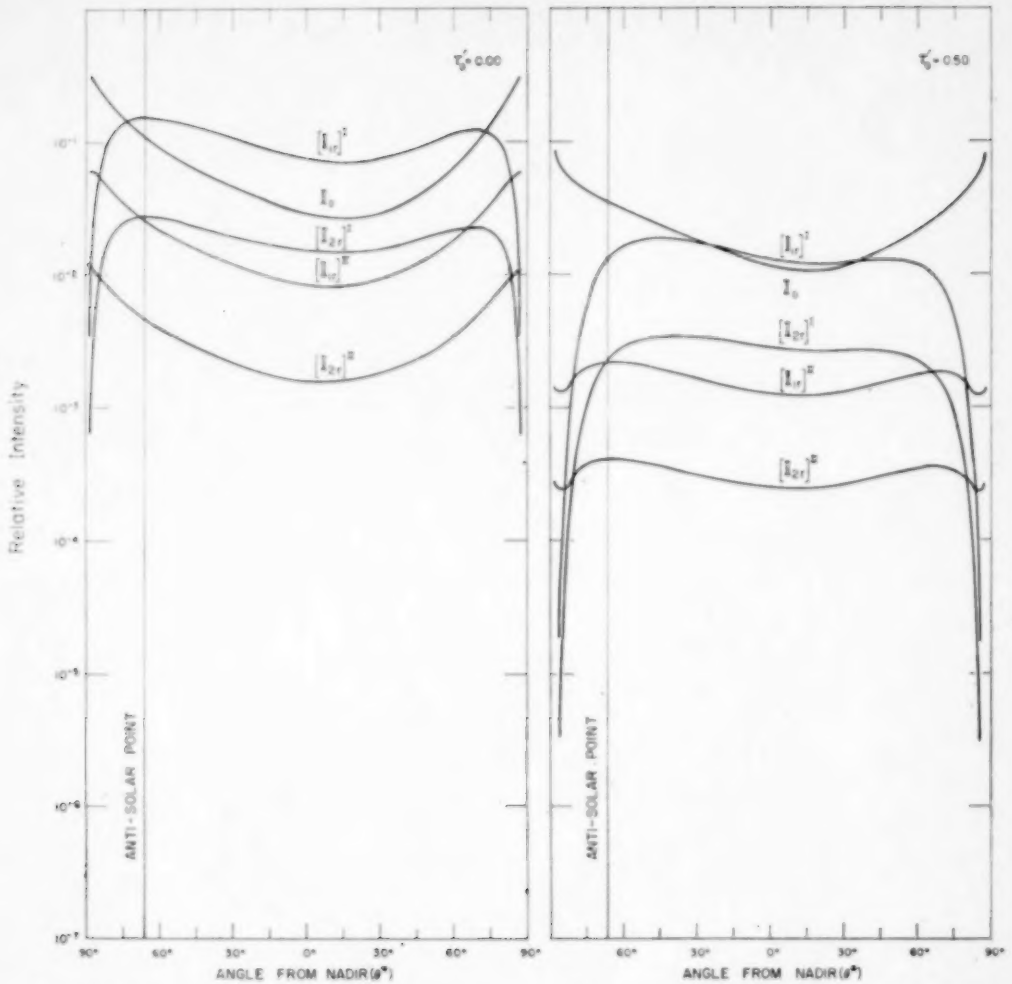


Fig. 4. Relative intensities of five different components of upward radiation for $\mu_0=0{\cdot}40$ ($\theta_0=66{\cdot}4°$).

the entire atmosphere, consists of five components each of which can be represented by a matrix. The intensity represented by matrix $\mathbf{I}_D$ gives the diffuse reflection of direct solar radiation $\pi\mathbf{F}_0(-\mu_0, \varphi_0)$ by the upper layer. The lower layer of scattering optical thickness, $\tau_1-\tau_0$ is being illuminated from above by (*a*) attenuated direct solar radiation $\pi\mathbf{F}_0 \exp(-\bar{\tau}_0/\mu_0)$ and (*b*) $\mathbf{I}^{(a)}(\bar{\tau}_0; \mu, \varphi; -\mu_0, \varphi_0)$ diffuse transmission of $\pi\mathbf{F}_0$ by the upper layer. Both of these radiations are diffusely reflected by the lower layer giving rise to illumination of the upper layer from below by radiations represented by matrices $\mathbf{I}^{(1R)}(\mu', \varphi')$ and $\mathbf{I}^{(2R)}(\mu', \varphi')$ respectively. Direct and diffuse transmission of $\mathbf{I}^{(iR)}(\mu', \varphi')$ by the upper layer is represented respectively by the matrices $[\mathbf{I}_{tr}]^1$ and $[\mathbf{I}_{tr}]^{11}$. In order to evaluate each of these components, we have solved the equation of radiative transfer for the upper layer when it is illuminated either from above or from below by a flux of parallel radiation. Since the lower layer has

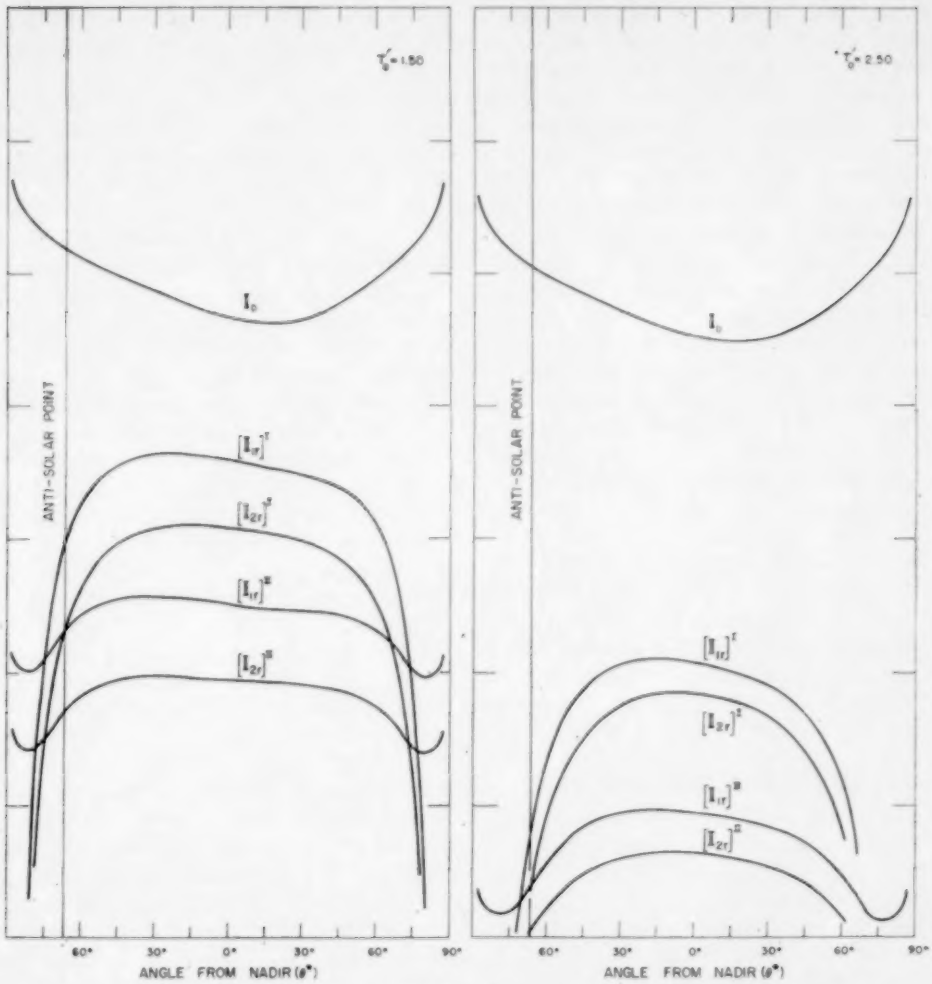


Fig. 4 cont'd.

no absorbing material, we have applied the solution derived by Chandrasekhar.

The scattering normal optical thickness $\tau_1 - \tau_0$ of the lower layer was taken to be 1·00 as the values of 12 scattering functions are available for this value of optical thickness. The scattering optical thickness (τ_0) of the upper layer was taken to be 0·18. Hence, the scattering optical thickness (τ_1) of the entire atmosphere is 1·18 corresponding to a wave length of 2990 Å. For an isothermal atmosphere of a scale height $H = 8$ km, the height corresponding to any value of τ can be obtained from the expression

$$\tau = \tau_1 e^{-h/H}. \qquad (2)$$

The height of the bottom of the upper layer as obtained from this relationship is 15·04 km. τ_m, the scattering optical thickness at the point of ozone maximum, was taken to be 0·06 which fixed the height of the ozone maximum at 23·84 km. For $\kappa = 2$ in equation (1) and the above mentioned values of τ_m and τ_0, 98 per

cent of the total ozone is in the upper layer. The ozone distribution represented by these constants can be considered fairly reasonable above 22 km, but below this level it gives ozone concentrations lower than the observed ones.

Since the aim of this study was to obtain a broad picture of the variations of diffusely reflected radiation under different conditions, the computations were restricted to only two azimuths $\varphi_0-\varphi^*=0°$ and 180°, i.e., to the sun's vertical. The direction of the plane of polarization is either perpendicular or parallel to the sun's vertical. For each of the five components discussed above, as well as for total radiation diffusely reflected by the entire atmosphere, the components parallel (I_l) and perpendicular (I_r) to the sun's vertical, the total intensity ($I=I_l+I_r$) and the degree of polarization (P) were computed for $F_0=1$ and for the following values of other parameters: $\tau'_0=0{\cdot}00$ (0·50) 2·50, $\mu_0=0{\cdot}98$, 0·92, 0·80, 0·60, 0·40, 0·28, 0·20, 0·10, 0·06 and 0·02, $\mu^*=1{\cdot}00$ (0·04) 0·04 and $\varphi_0-\varphi^*=0°$ and 180°.

All the computations were performed on the IBM Electronic Computer No. 709 of the Western Data Processing Center at the University of California, Los Angeles, California.

3. DISCUSSION OF THE RESULTS

The discussion of all the results of computations in detail would be very lengthy and unnecessary for many purposes. Hence, we limit our discussion to the total intensity in the case of each of the five components of the upward radiation for four different values of τ'_0, i.e., 0·00, 0·50, 1·50 and 2·50, and for the following five different values of μ_0: 0·98, 0·60, 0·40, 0·20 and 0·10. The discussion will cover all the directions of observations from one horizon to the other, through the anti-solar point and nadir.

The variation of the total intensity (I) for each of the five components of the upward radiation is shown for $\mu_0=0{\cdot}98$ (sun very near the zenith) and for four different absorption thicknesses in Fig. 2. In these diagrams the angle of observation (θ^*) from nadir is plotted as abscissa and the total intensity in logarithmic scale as ordinate. Figs. 3–6 give similar diagrams for the other four values of μ_0. The following points can be seen from these diagrams.

The variations in the magnitude of each of the five components relative to each other can be summarized in the following manner. When the sun is very near the zenith, $\pi\mathbf{F}_0\exp(-\tau_0/\mu_0)$ is always greater than the intensity of radiation $\mathbf{I}^{(s)}(\tau_0;-\mu,\varphi;-\mu_0,\varphi_0)$ integrated over the entire hemisphere. This is the consequence of the fact that in this case, the direct solar radiation passes practically along a vertical path, while the latter represents the diffuse transmission of the solar radiation by a layer of small optical thickness. Hence, for large values of μ_0, the intensity of the radiation defined by $\mathbf{I}_{1r}$ is always larger than that of the radiation given by $\mathbf{I}_{2r}$, for all values of τ'_0. With the increase of the zenith angle of the sun, attenuation of the direct solar radiation by the upper layer increases rapidly, and the intensity of $\mathbf{I}_{2r}$ becomes eventually equal or even larger than the intensity of $\mathbf{I}_{1r}$.

Because of the small optical thickness of the upper layer, the intensity of the component given by $\mathbf{I}_D$ is small compared to the intensity of the other component (Part I of $\mathbf{I}_{1r}$) when attenuation by the upper layer is small. With the increase of attenuation, the total radiation illuminating the top of the lower layer decreases rapidly and hence the intensity of $\mathbf{I}_D$ becomes greater than that of the radiation reflected by the lower layer. The value of the nadir angle θ^*, for which the intensity of $\mathbf{I}_D$ becomes one or two orders of magnitude greater than the intensity of radiation coming from the lower layer, can be easily estimated from the diagrams in Figs. 2–6 for different values of τ'_0 and μ_0.

4. EFFECT OF NEGLECTING HIGHER ORDERS OF SCATTERING IN THE UPPER LAYER

As stated above, only primary scattering in the upper layer was considered. This is bound to lead to an underestimation of the intensity of radiation diffusely reflected by this layer. The effect of multiple scattering in the upper layer in the presence of ozone can be evaluated by the method of numerical integration. However, these computations would be quite time

consuming and for many purposes, unnecessary. In the extreme case of $\tau'_0 = 0{\cdot}00$, i.e. when there is no ozone in the atmosphere, the tables of Coulson, Dave and Sekera[4] can be used for obtaining the intensity of radiation diffusely reflected by the atmosphere of scattering optical thickness 1·18, by extrapolation. This would give the intensity of radiation diffusely reflected when all the orders of scattering have been taken into consideration in the entire atmosphere. This extrapolated value has been compared with the previously computed intensity of radiation diffusely reflected by the atmosphere for the case $\tau'_0 = 0{\cdot}00$. The results are presented in the following table, where the ratio of the intensity of the radiation reflected by the atmosphere of the optical thickness $\tau_1 = 1{\cdot}18$ to the intensity of the radiation reflected by the atmosphere in which the multiple scattering is considered only in the lower layer, and only primary scattering in the upper layer of the optical thickness 0·18. Since the extrapolated values have been used in arriving at the figures given in the table, they are expected to be correct within $\pm 0{\cdot}02$.

Table I

μ^* \ μ_0	0·98	0·92	0·80	0·60	0·40	0·20	0·10
1·00	1·08	1·09	1·08	1·12	1·18	1·21	1·29
0·84	1·16	1·12	1·14	1·14	1·16	1·21	1·29
0·64	1·16	1·13	1·13	1·16	1·17	1·22	1·29
0·40	1·19	1·17	1·16	1·18	1·18	1·23	1·29
0·20	1·24	1·21	1·20	1·24	1·24	1·28	1·27
0·10	1·31	1·32	1·33	1·30	1·30	1·30	1·26

In this table it becomes evident that in the absence of ozone, multiple scattering in the upper layer can contribute from 8 to 15 per cent when μ_0 and μ^* are greater than 0·60, corresponding to solar zenith angles and the direction of observation from nadir up to 50°. For values of θ^* and θ_0 greater than 50°, the effect of multiple scattering increases rapidly from 15 to 30 per cent. However, under this condition, corrections due to the sphericity of the earth's atmosphere would become important too. It should be pointed out that the absorption due to the presence of the ozone would reduce the effect of multiple scattering in the upper layer.

5. CONCLUSION

The results of computations presented in this report are strictly true for a plane-parallel atmosphere of scattering optical thickness 1·18 which corresponds to only one wave length, namely 2990 Å at the sea-level. In this case, the curves presented in Figs. 2–6 stand for different total ozone amounts in the upper layer, $\tau'_0 = \alpha x_0$ where x_0 is the total ozone amount in cm at N.T.P. in the upper layer and α is the exponential absorption coefficient of one cm of ozone at N.T.P.

However, the absorption coefficient of ozone changes much more rapidly with wave length than the scattering optical thickness and therefore one can consider τ'_0 as the only parameter variable in the wave length. For the value of 0·25 cm of total ozone and the values of absorption coefficients given by Vigroux[5], the wave length corresponding to $\tau'_0 = 0{\cdot}50$, 1·00, 1·50 2·00 and 2·50 corresponds approximately to 3125, 3070, 3040, 3018 and 3000 Å respectively. The wave length corresponding to $\tau'_0 = 0{\cdot}00$ may be taken to be around 3300 Å where the absorption is very weak.

In the preceding treatment, we have assumed that the ground is a perfect absorber of all the radiation falling on it and that there are no aerosol particles or clouds in the lower layer. The presence of both will definitely increase the contribution from the lower layer. Coulson[6] has computed the effect of a ground reflecting according to Lambert's law*, on diffusely reflected radiation. His results show that for scattering optical thickness 1·00 and the sun in the zenith, the intensity of diffuse radiation from the nadir is increased by a factor of 1·4 and 2·6 for $\lambda_0 = 0{\cdot}25$ and 0·80 respectively. For solar zenith angles 66·4°, the ratios decrease slightly to 1·2 and 2·1 respectively.

Thus, for ground reflectivity of 80 per cent,

* Radiation reflected by the ground is a certain fraction (λ_0) of total incident flux; the reflected radiation is isotropic and unpolarized irrespective of the state of polarization of incident radiation.

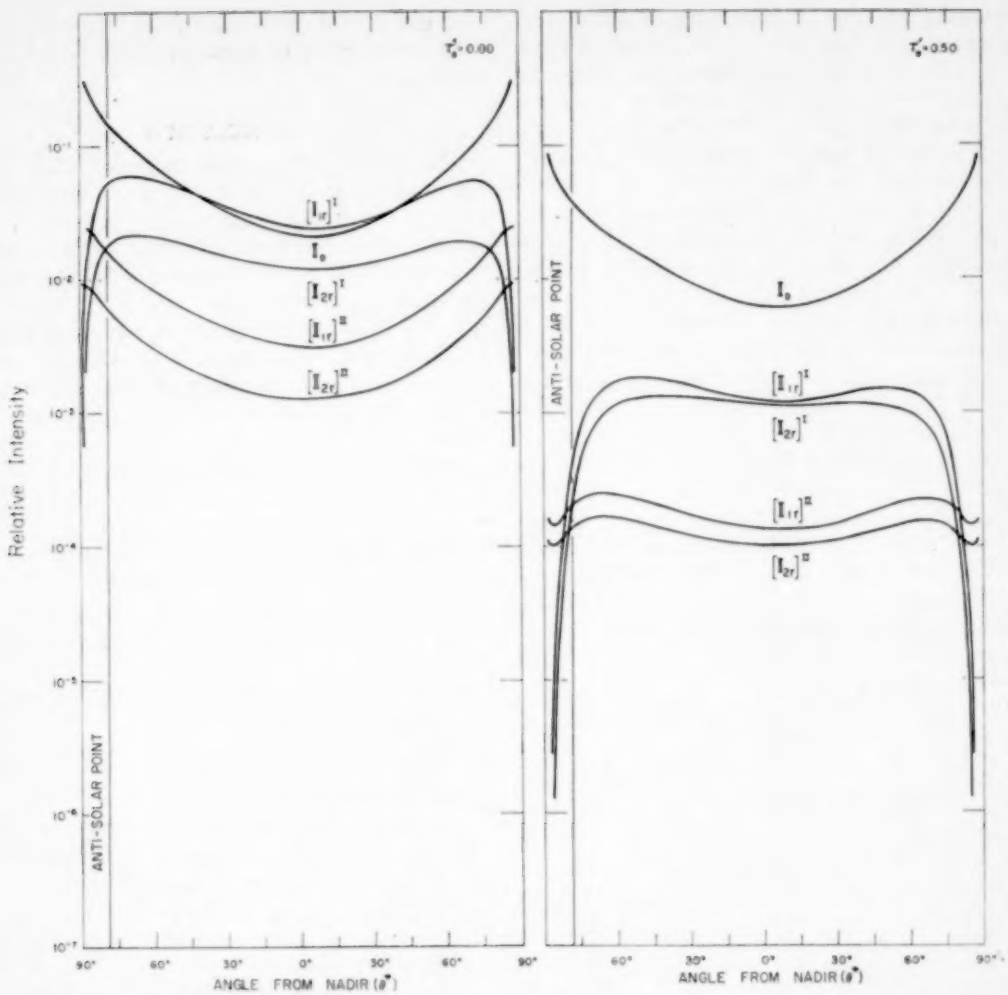


Fig. 5. Relative intensities of five different components of upward radiation for $\mu_0=0{\cdot}20$ ($\theta_0=78{\cdot}5°$).

i.e. ground covered with fresh snow, the contribution from the lower layer should be increased by a factor up to 3, the exact figure depending upon the direction of observation and of incident radiation. In the presence of clouds, the factor will also depend upon the height of the cloud above the ground and is likely to be as large as 10 for high level clouds.

Hence, under normal atmospheric conditions, the effect of the ground reflection or the presence of dust and haze, or eventually of low level clouds, may lead to the increase of the intensity of radiation emerging from the lower layer by a factor of 3 to 5. If we select, for the intensity measurements of the upward radiation, the limit of applicability by the requirement that the contribution from the lower layer should be less than 2 per cent, then the spectral region of wave length greater than 3070 Å is ruled out completely. It is interesting to note that this is the part of the spectrum generally used for the determination of the vertical distribution of

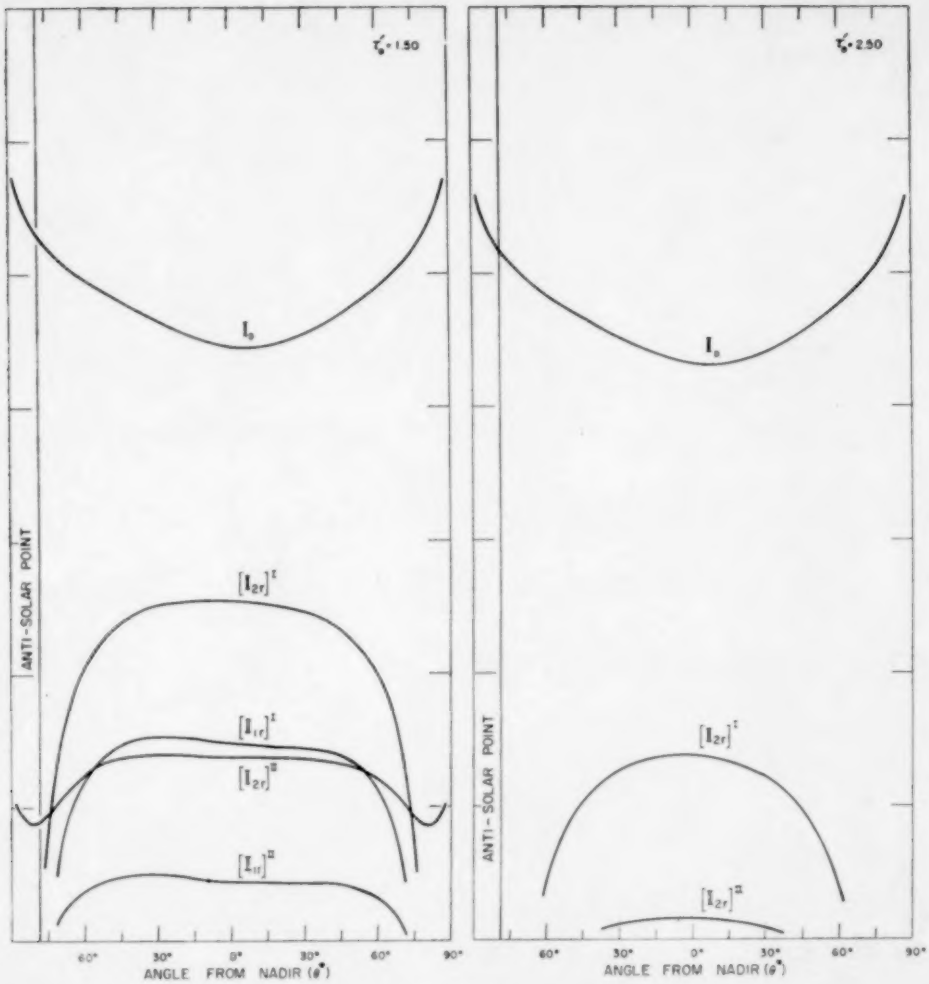


Fig. 5 cont'd.

ozone from the "umkehr" measurements. The spectral region between 3000 and 3070 Å can be used for the determination of ozone concentration of the lower stratosphere only for lower sun elevations and if the atmospheric conditions are fair. On the other hand, for high sun one would have to use wave lengths lower than 3000 Å. One can see that there is a great danger of losing ozone measurements in the region 15 to 24 km unless full advantage is taken of the region 3000–3070 Å by maintaining constant checks on the contribution from the lower layer. This can be done by the measurement of polarization of diffusely reflected radiation. When the contribution of the lower layer becomes considerable, it should be possible to compute its magnitude theoretically and apply proper corrections to the observed values. These corrections, though approximate in nature, would be of great help in the evaluation and understanding of the data.

The observations in the directions up to 60°

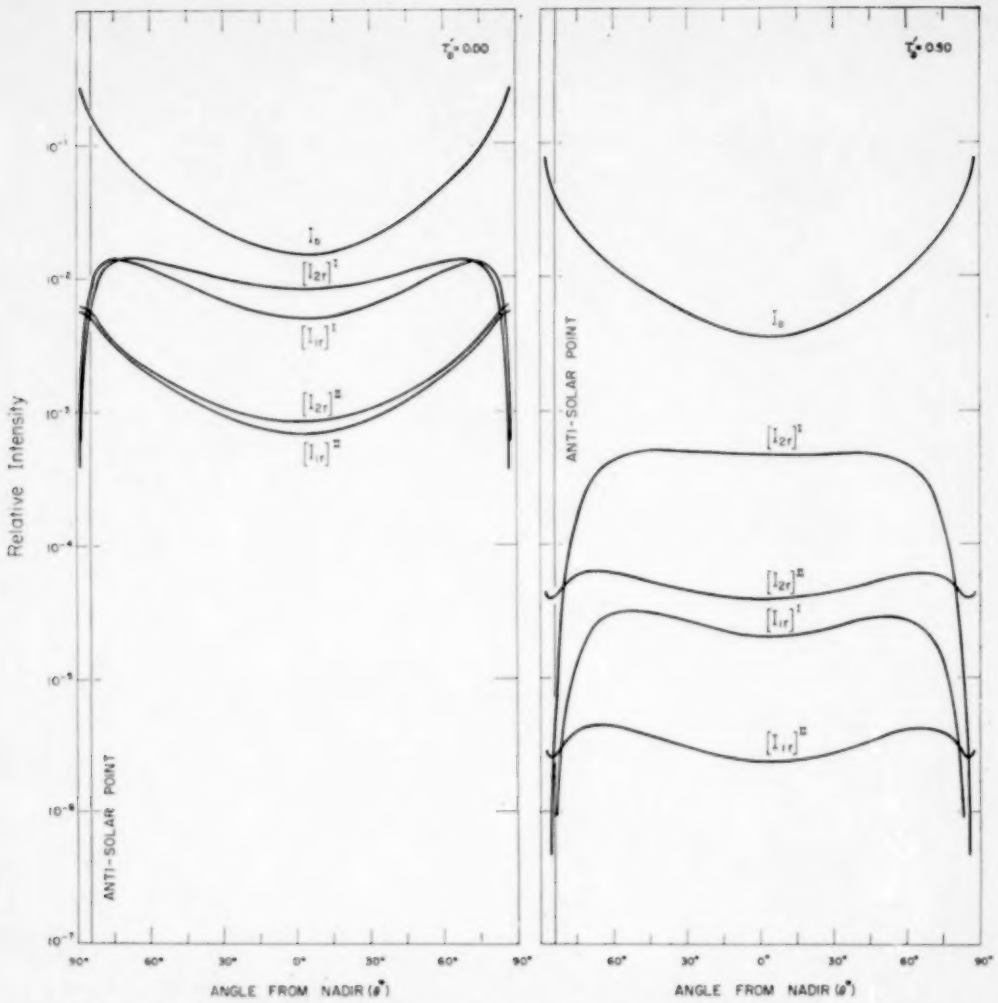


Fig. 6. Relative intensities of five different components of upward radiation for $\mu_0 = 0{\cdot}10$ ($\theta_0 = 84{\cdot}2°$).

from nadir will be definitely of some help in the evaluation of the data. The region of the spectrum below 2800 Å is not expected to be of any use unless one is interested in fine structure of ozone distribution above 50 km.

One can now estimate the contribution from the lower atmosphere which has been neglected by Singer and Wentworth. It may be mentioned that ozone distribution adopted by them is not identical with the one used by us. However, we can consider our results for $\tau'_0 = 2{\cdot}50$ corresponding to about 0·25 cm of ozone above 15 km for the wave length 3000 Å. This case is very near to a rather extreme case of 0·30 cm of ozone above 15 km taken by these authors. Our computations show that, in this case, the lower layer contributes as much as 19 per cent when the solar zenith angle is 11·5°. The contribution decreases to 17 per cent, 11 per cent, 4 per cent and 1 per cent for solar zenith angles of 23·1°, 36·9°, 53·1° and 66·4° respectively. As mentioned before, snow cover or

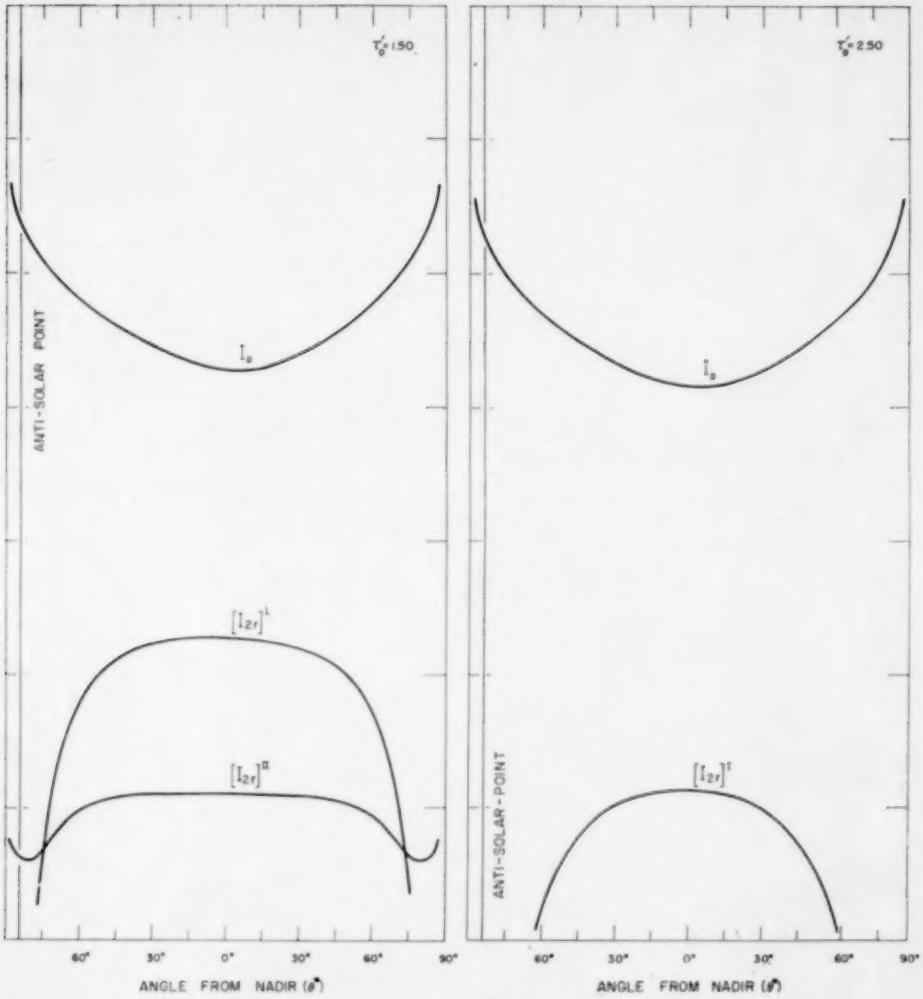


Fig. 6 cont'd.

clouds can increase these factors appreciably and thus annihilate all chances of obtaining reliable results.

The instrumentation problem for measuring ozone from satellites or high level balloons reduces to a construction of a reliable instrument capable of measuring radiation at about ten different wave lengths in the spectral region 2800–3070 Å. The permissible band-width of the instrument at any point of the spectrum would depend upon wave length and energy consideration but it would be around 10 Å. This would probably mean the use of high quality interference filters rather than a double monochromator.

The detailed specifications for exact wave length, band-width, angular aperture of the instrument and the number of observations in a scan, would require a thorough study of the particular problem at hand. Minute details of

computations and the exact figures omitted here may be of some assistance in special study and may be made available on request.

Acknowledgement—The machine time for the computation was made available, without charge, by the Western Data Processing Center, University of California, Los Angeles California. Thanks are extended to Mr. Kendall Wright, Chief of Machine Operations, Western Data Processing Center, for his personal interest and valuable advice at various stages of work. It is a pleasure to acknowledge with thanks the assistance given by Miss Elizabeth Jankus in writing the program for I.B.M. Electronic Computer 709 and in carrying out check computations on desk calculator. Finally, thanks are extended to Miss Avis St. John, Miss Bettina Klaus and Miss Geraldine Petersen for the assistance rendered by them in putting this report into its final form.

REFERENCES

1. S. F. SINGER and R. C. WENTWORTH, *J. Geophys. Res.*, **62**, 299 (1957).
2. Z. SEKERA and J. V. DAVE, *Diffuse Reflection of Solar Ultraviolet Radiation in the Presence of Ozone*. Scientific Report No. 4 under Contract No. AF 19(604)-2429 (1960).
3. S. CHANDRASEKHAR, (1950): *Radiative Transfer*. Clarendon Press, Oxford (1950).
4. K. L. COULSON, J. V. DAVE and Z. SEKERA, *Tables Related to Radiation emerging from a Planetary Atmosphere with Rayleigh Scattering*. University of California Press, Los Angeles (1960).
5. E. VIGROUX, *Contribution à l'étude expérimentale de l'absorption de l'ozone*. Thesis Sci. Faculty University of Paris, Ser A, No. 2542 (1953).
6. K. L. COULSON, *The Flux of Radiation from the Top of a Rayleigh Atmosphere*. Scientific Report No.1 under Contract No. AF 19(604)-2429 (1959).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 137–140. Printed in Great Britain.

THE NEAR-INFRA-RED SPECTRUM OF METEORS*

PETER M. MILLMAN and IAN HALLIDAY

National Research Council, Ottawa, Canada; Dominion Observatory, Ottawa, Canada

(*Received 8 August 1960*)

Abstract—The photographic infra-red spectra of nine meteors are studied. A total of 15 atomic emission lines are identified in the infra-red region, due to NI, OI, and CaII. As yet no infra-red lines have been observed in the spectra of slow meteors. The lines due to CaII become enhanced relative to lines of NI and OI with increasing luminosity of the meteor.

The first photographic records of the infra-red spectrum of meteors were secured by Millman[1] in 1950. In spite of the considerable effort which has been made on the Canadian meteor observing programmes to obtain meteor spectra in the photographic infra-red, further examples have accumulated rather slowly. At the end of 1959, nearly ten years after the first two infra-red spectra were photographed, only nine spectra showing lines in the infra-red had been obtained. Some negative results are available from a few other spectra, which, although photographed on infra-red film, failed to record any lines in the infra-red.

The nine spectra are listed in Table 1. The first column is the serial number of the spectrum in Millman's world list of meteor spectra[2]. The next column lists the date and Universal Time of the meteor followed by the shower membership (*P* for Perseid, *N* for non-shower). The next column contains the photographic magnitude at maximum light reduced to a standard distance of 100 km in the zenith. This has been derived from visual magnitude estimates. Heights above sea-level are listed in the last three columns.

The nine spectra were photographed at three stations, the Metcalfe Road Station and the Springhill Meteor Observatory, both near Ottawa, Ontario, and the Newbrook Meteor Observatory, Alberta. The spectrographs used had focal lengths ranging from 50 to 200 mm and were equipped with replica transmission gratings giving dispersions in the range 100 to 900 A/mm. The film used was in all cases a Kodak high speed infra-red emulsion. This has a high sensitivity out to approximately 9000 A.

Brief notes on the individual spectra follow:

Spectrum 112. This spectrum shows 5 features in the infra-red, identified with 2 oxygen and 3 nitrogen multiplets. Heights were determined by three station radar triangulation.

Spectrum 114. This is a very faint record, showing only one feature in the infra-red, the oxygen triplet at 7774 A. Heights were determined as for Spectrum 112.

Spectrum 138. Some 7 features could be measured in the infra-red region of this spectrum. Wavelengths and identifications are listed in Table 2. Reproductions of this spectrum appear in Ref. (1) and in Fig. 2 of this paper. Approximate heights have been determined by assuming a standard Perseid radiant position for this meteor.

Spectrum 224. The only feature seen in this spectrum, apart from the zero order image, is the oxygen triplet at 7774 A. No height or velocity information is available.

Spectrum 272. This brilliant, blue-green, meteor was observed at Springhill and estimated as of visual magnitude at least as bright as −7. It had a persistent train that remained visible for 3 min and the radar echo was recorded for over 6 min on a frequency of 32 mc/s. The meteor was photographed on three cameras, a very good spectrum of the wake being secured on a panchromatic plate. The infra-red plate shows the blue spectrum out to the 20th order and the infra-red spectrum to the 9th order. The infra-red multiplet, CaII (2) was

* Issued as Contribution from the Radio and Electrical Engineering Division, National Research Council, No. 6030, and Contributions from the Dominion Observatory, Vol. 4, No. 8.

Table 1. Infra-red Meteor Spectra

No.	Date	U.T. h.	U.T. m.	U.T. s.	Shower	M_{pm}	Heights in km b	Heights in km max	Heights in km e
112	1950 Aug. 12	6	29	43	P	–5	110	(90)	85
114	13	6	58	05	P	–5	117		90
138	1953 Aug. 14	7	20	13	P	–4	114	97	94
224	1957 Aug. 14	8	58	37	N				
272	1958 Aug. 12	6	20	15	P	–10	112	90	83
273	12	6	39	25	P	–7	104	84	79
275	12	8	01	15	P	–9	106	79	79
(U740)	1959 Aug. 12	7	15	*	P				
(H1401)	13	4	55	*	P		112	108	106

recorded in the wake, particularly in the low orders. Portions of the 5th and 6th orders in the infra-red are reproduced as the third spectrum in Fig. 2. Heights for this meteor were determined by combining the photographic positions with the radar ranges.

Spectrum 273. This spectrum, secured at Springhill, is reproduced in Fig. 1, while the lines measured appear in Table 2. Heights were determined in the same manner as for Spectrum 272.

Spectrum 275. This spectacular Perseid was photographed by seven cameras at the Meanook and Newbrook observatories. Some of the panchromatic films show great detail but the only infra-red record fell at the edge of the camera field. As a result only one infra-red multiplet was recorded, as shown in Table 2. Heights were determined by a triangulation of photographs from the two stations.

Spectrum (U740). Only lines of ionized calcium were recorded in this spectrum. Eight lines, appearing in a short burst, are identified with the second and third orders of CaII (2) near 8600 A, and the sixth order of the *H* and *K* lines, all on the side of the zero order remote from the blazed orders.

Spectrum (H1401). The major burst of this meteor is shown in the middle spectrum of Fig. 2 and measured wavelengths are given in Table 2. Heights were determined in the same manner as for Spectrum 138.

A total of seven multiplets have been identified in the infra-red meteor spectra, as shown in Table 2, four due to atomic nitrogen, two due to atomic oxygen, and one of ionized calcium. The six multiplets of nitrogen and oxygen are all of quite high excitation, ranging from 10·7 to 12·1 eV. The calcium multiplet, from an ionized atom, would also be considered a high-excitation feature, although the excitation energy, after ionization has been achieved, is only 3·1 eV.

Such high-excitation multiplets are generally found in the spectra of fast meteors but not in slow meteors. One might then expect that the spectra of slow meteors would show few, if any, strong lines in the infra-red. At present there is little observational evidence on this point. One rather faint Geminid, Spectrum 315 with velocity about 35 km/sec, was photographed on infra-red film and showed no infra-red lines. Spectrum 210, an asteroidal-type meteor with a very low velocity of only 13 km/sec has been studied by Halliday[3]. This was photographed on infra-red film without showing any spectral lines in this region. Spectrum 225 also failed to record any infra-red lines. The photographic trail is long, over 3 sec in duration, and the spectrum in the visual region suggests that this was a slow meteor. All of these spectra were weak, so it remains possible that a very bright, slow meteor might have an interesting spectrum in the infra-red. At present, however, it seems that strong infra-red lines are correlated with high meteoric velocities.

From Table 1 we see that the meteors discussed here appeared on the average at a

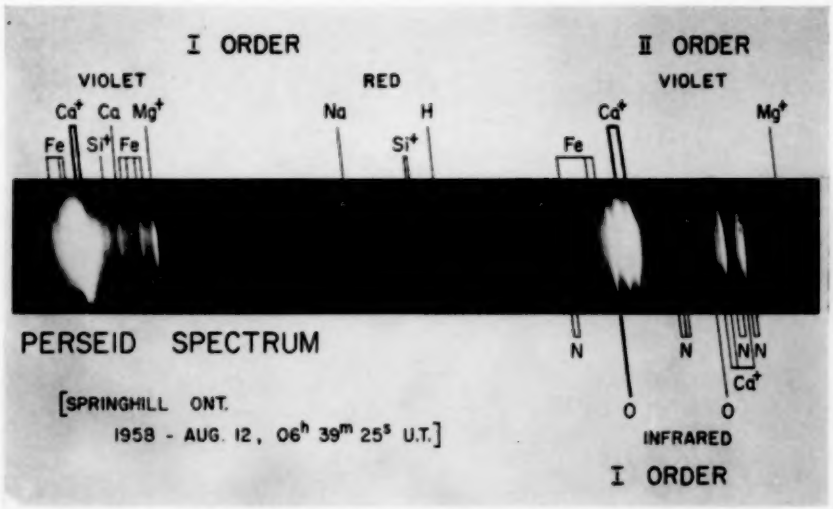


Fig. 1. Spectrum 273, showing the strongest segment. It will be noted that the first order infra-red spectrum is overlapped by the second order violet.

(facing page 138)

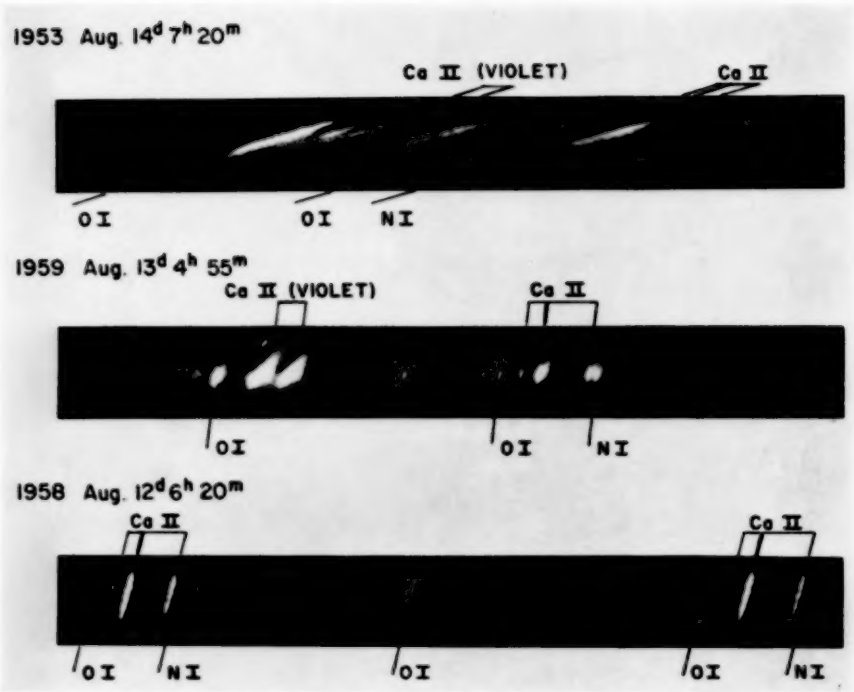


Fig. 2. Three infra-red Perseid spectra arranged to illustrate changes in relative intensity between CaII and OI and NI. The CaII lines from the violet overlap two of these infra-red spectra.

Table 2. Measured Wavelength and Identification

Spectrum 138	Spectrum 273	Spectrum 275	Spectrum (H 1401)	Multiplet* identification
			3724·0	FeI (5) (21)
3748	3745		3738·3	FeI (5) (21)
	3761			FeI (21)
	3820·4			FeI (4) (20)
			3829·8	MgI (3)
3840	3836·3		3838·8	MgI (3) FeI (20)
	3857·0		3859·7	FeI (4)
3932	3933·0		3933·7	CaII (1)
3967	3967·9		3969·0	CaII (1)
	4003·8		4004·3	FeI (43)
			4034·0	MnI (2)
4051	4046·5		4045·4	FeI (43)
	4064·6			FeI (43)
	4074·0			FeI (43)
	4102·0			SiI (2) FeI (18)
4130·4	4131·4			SiII (3)
4227·0	4227·2		4226·3	CaI (2)
	4256·7			FeI (42) (152)
4267·7	4271·6			FeI (42) (152)
4303·3	4307·1			FeI (42)
	4327·0			FeI (42)
	4352·1			MgI (14)
4383·2	4383·2			FeI (41) (2)
	4404·9			FeI (41)
4427·3	4429·6			FeI (2)
	4461·8			FeI (2)
4481·1	4482·8		4481·2	MgII (4) FeI (2)
	4527·1			FeII (37) (38)
	4554·7			FeII (37) (38)
4573·6	4581·7			FeII (37) (38)
5892·5	5892·2			NaI (1)
6347·2	6347·4			SiII (2)
	6371·3			SiII (2)
	6569·3			HI (1)
		7423·7		NI (3)
7450·8	7443·6	7442·1	7443·0	NI (3)
	7467·4	7468·0	7469·4	NI (3)
7770·2	7773·6		7772·7	OI (1)
8188·4	8186·8		8185·2	NI (2)
8218·0	8217·3		8216·6	NI (2)
	8243·6		8242·2	NI (2)
8444·1	8445·4		8446·9	OI (4)
	8499·1		8496·7	CaII (2)
8537·6	8542·2		8542·2	CaII (2)
	8596·2		8598·2	NI (8)
	8630·0		8630·3	NI (8)
	8661·8		8662·8	CaII (2)
8686·0	8683·0		8682·7	NI (1)
	8715·6		8714·1	NI (1)

* Identification numbers are from *A Multiplet Table of Astrophysical Interest* by C. E. Moore, Princeton Observatory Publications, No. 20 (1945).

height near 110 km, while maximum light was at 90 km, and the point of disappearance a few km lower. These are normal heights for bright Perseid meteors. For most of the meteors listed in Table 1 the oxygen triplet at 7774 A is the first feature to appear in the spectrum. Towards the end of the trail the oxygen generally loses strength relative to the ionized calcium. This is particularly true in the case of terminal bursts.

Although all the infra-red spectra are similar in appearance there are marked differences in the relative intensities of certain multiplets. Probably the most outstanding example of this is illustrated by the three spectra reproduced in Fig. 2. These have been selected to demonstrate the relative intensities between the CaII triplet near 8600 A and the two multiplets on either side, OI at 8446 A and NI near 8700 A. In the 1953 spectrum OI and NI are very strong while CaII can barely be seen between them. In the 1959 example all the multiplets are of comparable intensity. In the 1958 spectrum CaII is very strong while OI and NI can barely be detected in the orders reproduced. This variation seems to be correlated with absolute luminosity as the three spectra are arranged in order of increasing luminosity. This also agrees with the behaviour of OI (1) mentioned earlier. It may be a little surprising that this line with a fairly high excitation appears so strongly and so early in fast meteor spectra. The explanation for such behaviour probably lies in the importance of atomic oxygen at higher levels in our atmosphere.

Further observational material is required before our knowledge of the photographic infra-red region of meteor spectra will compare with our knowledge of the visual region. Observations of sufficiently bright meteors in the velocity ranges below 60 km/sec (the velocity of Perseid meteors) are required. A significant improvement would be achieved by the use of filters to eliminate overlapping orders. At present numerous lines from the blue region of the second-order spectrum are intermingled with the first-order infra-red lines. Some confusion in line identifications may result. With this overlapping eliminated the better spectra would then be more suitable for a detailed study which would include photometric analysis.

REFERENCES

1. P. M. Millman, *Nature* **172,** 853 (1953).
2. P. M. Millman, *J. R. Astr. Soc. Can.* **53,** 271 (1959).
3. I. Halliday, *Astrophys. J.* **132,** 482 (1960).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 141–152. Printed in Great Britain.

TEMPERATURE AND CIRCULATION OF THE VENUS ATMOSPHERE*

YALE MINTZ
Department of Meteorology, University of California, Los Angeles

(*Received* 21 *September* 1959; *Revised* 11 *July* 1960)

Abstract—The approximate radiative equilibrium temperature is computed for the atmosphere above the visible surface of Venus, taking carbon dioxide as the only strongly absorbing gas on the planet. It is found that with radiative equilibrium the atmosphere is convectively stable, and that the mean temperature of the visible surface is 237°K.

The small difference in observed intensity of infra-red radiation from the day and night sides of the planet is attributed to a large heat capacity of the emitting body, and from this it is inferred that the visible surface is only a thin layer of smog covering a universal (but non-aqueous) ocean.

The appearance of the clouds of Venus in visible and in ultra-violet light suggests a form of the general circulation in which there are large low-level horizontal cells that transport heat from low to high latitudes and from the day to the night side of the planet, surmounted by a circumpolar vortex.

1. RADIATION BUDGET AND TEMPERATURE OF THE ATMOSPHERE ABOVE THE VISIBLE SURFACE

The only gas that has been positively identified in the atmosphere of Venus is carbon dioxide.† From the intensity of the absorption lines in the reflected sunlight, Herzberg[1] estimated that the amount of CO_2 above the visible surface of Venus was of the order of 1000 meter-atmospheres (at standard temperature and pressure). For our model, we shall adopt this as the amount of CO_2 above the visible surface of the planet. We assume that if other gases are present they are, by comparison with the CO_2, of negligible importance for the gross radiative properties of the atmosphere above the visible surface.

A second assumption we shall make is that the visible surface radiates as a black body in the infra-red, beyond say 3 microns.

Infra-red absorption by CO_2 is such that, as a rough approximation, a layer of the order of 100 m-atm (STP) will act as the equivalent of a perfect absorber in the wavelengths 3 to 5·5μ and 12·5 to 17·5μ (cf. Callender[2], Ohring[3], Howard, Burch and Williams[4]). CO_2 also absorbs in narrow bands at 9·4 and 10·4μ, and again, as a rough approximation, we will treat these absorptions by 100 m-atm of CO_2 as the equivalent of a one micron wide band of complete absorption centered at 10μ. At wavelengths longer than 17·5μ there is no known absorption by CO_2. Below 3μ the absorption is of little importance in our problem because of the weak radiation in these wavelengths at the temperatures we are concerned with.

In our model, then, a 100 m-atm isothermal layer of CO_2 will be a black body in the wavelengths 3·0 to 5·5μ, 9·5 to 10·5μ and 12·5 to 17·5μ; and it will be completely transparent in all other wavelengths. Inasmuch as we do not know the exact amount of CO_2 above the visible

* A report on research carried out at The RAND Corporation, Santa Monica, as part of the U.S. Air Force sponsored program of the Planetary Sciences Group.

† Since this paper was submitted for publication it has been reported by J. Strong that spectroscopic observation from a high level balloon shows something like 0·02 mm of precipitable water vapor above the visible surface of Venus. Infra-red absorption by this amount of water vapor, which is less than one-tenth of one per cent of the total water vapor above the ground in our own atmosphere, is too small to affect the gross features of the radiation budget of Venus as described in this paper.

surface of Venus, but only its order of magnitude, this simplified absorption spectrum for CO_2 will suffice for our computations.

The radiation from a 100 m-atm isothermal layer of CO_2 as a function of temperature, when using our simplified absorption spectrum, is given in Table 1.

Table 1.

Temp (°K)	160	180	200	220	240	260	280
Radiation (cal cm^{-2} min^{-1})	0·004	0·017	0·034	0·053	0·080	0·115	0·160

With this absorption spectrum we can compute the radiation budget and the temperatures of the visible surface and the overlying atmosphere by dividing the atmosphere into ten isothermal layers of equal mass, for which the equations for radiative equilibrium are

$$(1)\quad -2Q_1 + Q_2 \ldots\ldots\ldots\ldots\ldots\ldots = 0$$
$$(2)\quad Q_1 - 2Q_2 + Q_3 \ldots\ldots\ldots\ldots\ldots = 0$$
$$(3)\quad Q_2 - 2Q_3 + Q_4 \ldots\ldots\ldots\ldots = 0$$
$$\vdots$$
$$(9)\quad Q_8 - 2Q_9 + Q_{10} \ldots\ldots = 0$$
$$(10)\quad Q_9 - 2Q_{10} + Q_{g,\mathrm{co}_2} = 0,$$

where Q_1 through Q_{10} are the radiations emitted by the respective atmospheric layers, and Q_{g,co_2} is the radiation emitted by the visible surface that the lowermost atmospheric layer absorbs. All the atmospheric layers radiate in two directions; but the top layer, unlike the others, receives radiation only from below. The CO_2 absorption of solar radiation is very small and is neglected in this model.

In addition, we have the equation for radiative equilibrium of the visible layer,

$$(11)\quad S(1-a) + Q_{10} - Q_{g,\mathrm{co}_2} - Q_{g,\mathrm{window}} = 0,$$

where S is the solar radiation that falls on the visible surface, a is the integrated albedo of the visible surface (the fraction of the solar radiation that the surface reflects), and $Q_{g,\mathrm{window}}$ is the radiation emitted by the visible surface in the wavelengths which the overlying CO_2 atmosphere does not absorb.

Q_{g,co_2} and $Q_{g,\mathrm{window}}$ are both known functions of temperature when we assume that the visible layer radiates as a black body in the wavelengths greater than 3μ. Consequently, the ratio of the two is a known function of Q_{g,co_2}, or

$$(12)\quad \frac{Q_{g,\mathrm{co}_2}}{Q_{g,\mathrm{window}}} = E(Q_{g,\mathrm{co}_2}).$$

Given S, the incident solar radiation, and a, the integrated albedo, we can solve equations 1 through 12 for the twelve unknowns, $Q_1 \ldots Q_{10}$, Q_{g,co_2} and $Q_{g,\mathrm{window}}$, and from these find the corresponding radiative equilibrium temperatures of the atmospheric layers and the visible surface.

For S we will use the average incident solar radiation, 0·958 cal cm^{-2} min^{-1}, which is one-fourth the solar constant for Venus; and for a we will use 0·80, the value of the integrated albedo determined by Kozyrev[5].

With these values for S and a we obtain, from equations 1 through 12, the radiations, Q, shown in the second column of Table 2; and, by interpolation in Table 1, the corresponding temperature of each layer, shown in the third column of Table 2.

In addition, we obtain from these equations, $Q_g = (Q_{g,\mathrm{co}_2} + Q_{g,\mathrm{window}}) = (0{\cdot}077 + 0{\cdot}185)$ cal cm^{-2} min^{-1}; and the corresponding radiative equilibrium temperature of the visible surface, $T_g = 237°\mathrm{K}$.

Table 2 refers to the mean radiations and temperatures for the planet as a whole. But if we repeat the calculations for the equator of the planet, where the mean incident solar radiation, S, is the solar constant divided by π, we obtain: $Q_1 = 0{\cdot}009$, $Q_g = (Q_{\mathrm{co}_2} + Q_{g,\mathrm{window}}) = (0{\cdot}103 + 0{\cdot}236)$ cal cm^{-2} min^{-1}; and the corresponding equilibrium temperatures, $T_1 = 169°$ and $T_g = 253°\mathrm{K}$.

The equilibrium mean radiation budget is summarized in Fig. 1. The incident solar radiation penetrates the CO_2 atmosphere without absorption and falls on the visible surface,

Table 2.

Layer	Q	T	P_{top}— P_{bottom}	Z_{top}— Z_{bottom}
1	0·007	166°K	0 — 16·5 mb	∞ — 10·07 km
2	0·014	176	16·5 — 33·0	10·07 — 7·33
3	0·021	185	33·0 — 49·5	7·33 — 5·65
4	0·028	194	49·5 — 66·0	5·65 — 4·41
5	0·035	202	66·0 — 82·5	4·41 — 3·39
6	0·042	209	82·5 — 99·0	3·39 — 2·54
7	0·049	216	99·0 — 115·5	2·54 — 1·80
8	0·056	222	115·5 — 132·0	1·80 — 1·15
9	0·063	228	132·0 — 148·5	1·15 — 0·54
10	0·070	233	148·5 — 165·0 mb	0·54 — 0·00 km

where part is reflected and part penetrates into the visible layer. It is worth noting that the larger albedo of Venus compared with the Earth, (0·80 as against 0·35), more than offsets the greater solar constant for Venus (3·83 as against 2·00 cal cm^{-2} min^{-1}). As a result the effective solar radiation for Venus, $S(1-a)$, is 40 per cent *smaller* than for the Earth. This fact, taken alone, would lead one to expect characteristically lower temperatures on Venus than on the Earth.

Of the long-wave radiation from the visible surface about two-thirds passes through the overlying CO_2 atmosphere without absorption, with more than half of this going through the "window" beyond 17·5 microns. The remaining part of the surface radiation is absorbed by the overlying CO_2 atmosphere, but roughly ninety per cent of this energy is returned to the surface again as the atmospheric back radiation which produces the greenhouse effect. Without atmospheric absorption and the greenhouse effect, the equilibrium surface temperature, for an effective solar radiation of 0·192 cal cm^{-2} min^{-1}, would be only 220°K.

Not shown in the figure, and neglected in our model, is the small conductive–convective transfer of heat from visible surface to atmosphere that would arise due to the small temperature discontinuity at the boundary if the system is in radiative equilibrium.

Using the hydrostatic equation, we obtain the geometrical thickness of each 100 m-atm CO_2 layer,

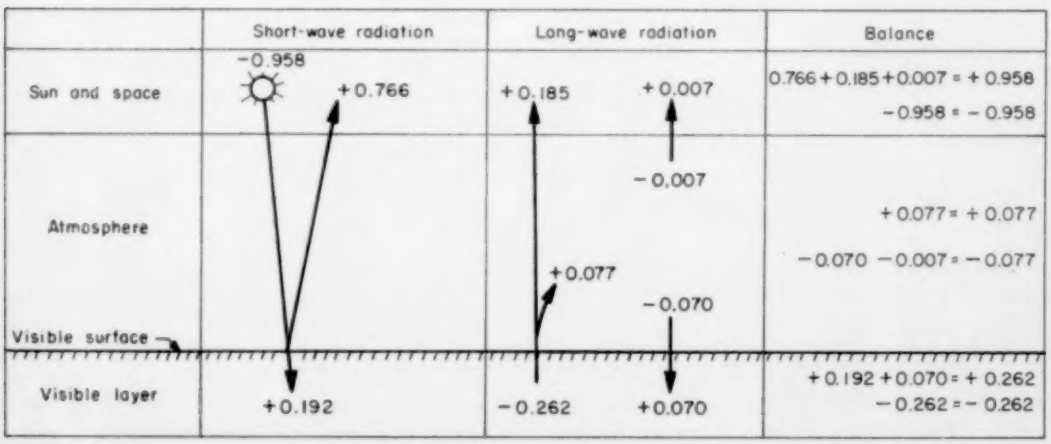


Fig. 1. Radiation budget of Venus.

$$\Delta Z = \frac{RT}{g} \ln\left(\frac{P_{\text{bottom}}}{P_{\text{top}}}\right),$$

where $R = 189\ kj\ t^{-1}\ \text{deg}^{-1}$ is the gas constant for CO_2, and $g = 8{\cdot}42\ \text{m}\ s^{-2}$ is the acceleration of gravity at the surface of Venus. The heights of the top and bottom of each isothermal layer are shown in the last column of Table 2.

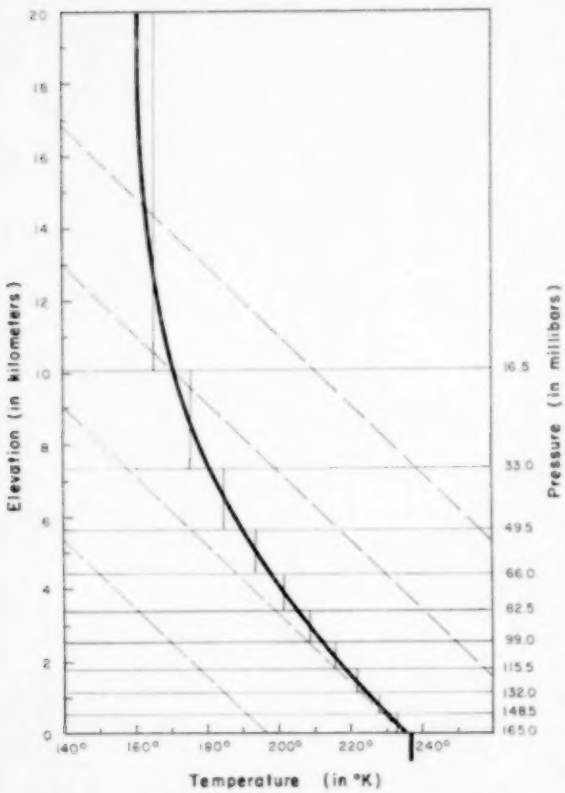


Fig. 2. Radiative equilibrium temperature of the atmosphere above the visible surface. (The thin broken lines are adiabats).

The radiative equilibrium temperature for each isothermal layer above the visible surface is shown by the vertical lines in Fig. 2; and the continuous line drawn through the centers of the isothermal layers is intended to represent the approximate temperature distribution for radiative equilibrium when the temperature varies continuously.

The thin broken lines in Fig. 2 show the adiabats for this atmosphere from

$$\frac{dT}{dZ} = -\frac{g}{C_p} = -10{\cdot}5°\text{C/km},$$

where $C_p = 800\ kj\ t^{-1}\ \text{deg}^{-1}$ is the approximate specific heat at constant pressure for CO_2.

The important feature shown in Fig. 2 is that with radiative equilibrium the temperature lapse rate is smaller than the adiabatic rate, and hence this atmosphere is convectively stable. In this important respect our model of the atmosphere above the visible surface of Venus differs from the Earth's atmosphere.

The reason why there is a convectively unstable lapse rate in the Earth's atmosphere is that on the Earth water vapor is the principal absorber of the long-wave ground radiation. The dependence of water vapor saturation pressure on temperature causes the water vapor to be concentrated in the lower levels of the atmosphere and, as a result, the radiative equilibrium lapse rate is especially large in the lower levels and exceeds the convective limit. But in our model of the Venus atmosphere the absorbing medium is the carbon dioxide itself, and the absorbing gas is uniformly distributed with height. As a result, the radiative equilibrium temperature gradient is not as great in the lower levels, and the atmosphere is stable.

The conclusion that the atmosphere above the visible surface of Venus is convectively stable depends, however, on the exact amount of CO_2 present. A smaller amount of CO_2 than 1000 m-atm would produce an even more stable lapse rate, and a larger amount would produce a less stable one. But inasmuch as the larger part of the outgoing long wave radiation comes from the visible surface, increasing or decreasing the amount of CO_2 above the visible surface would have little effect on the temperature of the visible surface. It would only be the upper levels of the CO_2 atmosphere that would have a lower or a higher temperature with the greater or lesser amount of CO_2.

Adding nitrogen to the carbon dioxide in our model of the Venus atmosphere would not affect the calculated temperature of the visible surface nor the temperature of each of the layers con-

taining 100 m-atm of CO_2. Adding nitrogen would increase both R and C_p, which would reduce both the radiative equilibrium lapse rate and the adiabatic lapse rate; but the change in the latter would not be as large and, as a result, the atmosphere with nitrogen added would be even more stably stratified under radiative equilibrium (it would suppress convection even more) than the pure CO_2 atmosphere.

What will to a sizeable extent affect the computed temperature of the visible surface is the value that is used for the integrated albedo. If we should use, for example, $a=0{\cdot}60$ (which is one of the older determinations of the visible albedo, by Russell[6]), this would double the mean effective insolation. When S is again taken as one-fourth of the solar constant, and $a=0{\cdot}60$, the solution of equations 1 through 12 gives: $Q_1=0{\cdot}016$, $Q_g=(Q_{g,\,CO_2}+Q_{g,\,\text{window}})=(0{\cdot}170+0{\cdot}368)$ cal cm^{-2} min^{-1}; and the corresponding radiative equilibrium temperatures, $T_1=178°$ and $T_g=284°$K.

2. COMPARISON WITH SOME OBSERVATIONS

The most careful observations of the infra-red radiation from Venus are those reported by Sinton[7] and Sinton and Strong[8]. They observed the radiation from Venus within the 8 to 13 micron band, corrected for absorption by the Earth's atmosphere, and reduced the measurements to equivalent black body temperatures both as a function of position on the planetary disk and integrated over the disk as a whole.

From their measurements over the whole disk of the planet they obtained the equivalent temperature—of a black body of the same size as the planet and yielding the same amount of radiation in the 8–13 micron band as does the planet—of 225°K. We may note here that an infra-red black body at this temperature would be in radiative equilibrium in the orbit of Venus if it absorbed 22 per cent of the Sun's radiation and reflected 78 per cent; and that this is in good agreement with Kozyrev's direct determination of the integrated albedo of Venus, 0·80.

According to our model, the radiation from Venus in the 8 to 13 micron band is a composite radiation, with the larger part of the radiation in this band coming from the visible surface and only a small part from the top of its CO_2 atmosphere. In our model, as shown in Table 2 and Fig. 2, the 225° composite temperature would correspond to the actual temperature at a fairly low level in the CO_2 atmosphere.

In their scans perpendicular and parallel to the terminator, which they refer to as equatorial and meridional scans, Sinton and Strong observed a falling off of the radiation toward the edges of the disk, similar to the limb darkening of the Sun. This limb darkening indicates that at least part of the 8 to 13 micron radiation from Venus comes from a region of the atmosphere in which the temperature decreases with height, which is consistent with our theoretical temperature curve.

A third important observation by Sinton and Strong is that the meridional scans, after correction for limb darkening, show a temperature decrease from the equator to the poles of Venus of some 20° to 30° centigrade. These scans give a temperature at the equator of 234°K, which is 9° centigrade higher than their mean temperature of 225°K for the planet as a whole. Inasmuch as these are both composite temperatures, we can regard their 9° excess of equatorial temperature over the mean temperature as consistent with our computations, which gave an equatorial excess of 16° at the visible surface and a 3° excess in the highest atmospheric layer; and the agreement is especially good if we recall that our computation was for radiative equilibrium at the equator, whereas on the true planet an atmospheric circulation of some kind will be produced by an equatorial temperature excess, and this will act to reduce the temperature excess.

A fourth important observation was obtained from their equatorial scans by Sinton and Strong, which confirmed the earlier work by Pettit and Nicholson[9] and others. The scans showed that the night side of Venus is only two or three degrees colder than the day side. Inasmuch as most of the radiation in the 8 to 13 micron band comes from the visible surface (in our model), this constancy of the day and night temperatures can give us some information about the physical character of the visible

surface. We shall discuss the implication of this observation further in the next section.

Another important observational study of Venus radiation is that of Chamberlain and Kuiper[10], who measured the relative intensities of the CO_2 rotational absorption bands near 0·8μ and obtained a mean temperature of 285° ±9°K. Kuiper[11] assigned the temperature found in this way to a level deep in the Venus atmosphere and relatively close to the visible surface. But this temperature is not consistent with our model when we use the albedo value of 0·80. Consistency would be obtained if we were to use a lower value of the albedo. As indicated earlier, using an albedo of 0·60 in our model yields a mean surface temperature of 284°K (and a still higher surface temperature at the equator). But the writer does not believe that fudging with the albedo to obtain this agreement is justifiable. He prefers to accept, tentatively, the lack of agreement as a weakness of his model.

Finally, among the direct observations, we must consider the measurements of centimeter wavelength radiation from Venus. But as this radiation, if it is of thermal origin, would come from below the visible surface, we will postpone its discussion until a later section.

3. THE CLOUDS IN THE ATMOSPHERE ABOVE THE VISIBLE SURFACE

The temperatures we derived for the levels high above the visible surface are not far from the sublimation temperature for carbon dioxide. At pressures of 10 to 30 mb the sublimation temperature for CO_2 is about 150° to 160°K, or some 10° to 20° lower than our computed radiative equilibrium temperatures at these levels. But if there are vertical motions at high levels in the Venus atmosphere, with total vertical displacements of the order of 1 to 2 km, adiabatic cooling would bring the carbon dioxide to its sublimation point. Cirriform clouds of CO_2 crystals might then form in the ascending air currents, with clear spaces in between where the air descends and is warmed adiabatically.

One would expect that a thin CO_2 cirrus cloud formed in this way would be difficult to see, just as the thin ice-cirrus clouds in our own atmosphere are difficult to observe and sometimes are revealed only where they form a halo around sun or moon. If observed at all, the CO_2 cirrus clouds of Venus ought to be revealed best by scattering of the Sun's ultra-violet light, and it is probably these clouds, in the form of bright patches and bands, that are seen in the ultra-violet photographs by Ross[12] and others.

One would not expect a thin cirrus cloud to have a large effect on the outgoing long-wave radiation. But the cloud might absorb enough of the infra-red radiation from the visible surface to produce an observable effect. In this connection it is noteworthy that Sinton and Strong found that a cold region that appeared at the north cusp, in two of their radiometric meridional scans, was in the same position as a bright cloud seen in ultra-violet photographs taken on the same dates.

4. THE VISIBLE SURFACE AND THE LAYER BELOW IT

We have, so far, assumed for the visible surface only that it is a black body radiator in the infra-red beyond 3 microns. By setting the radiation from this surface equal to the solar energy absorbed plus the atmospheric back radiation, we derived the mean temperature of the visible surface of 237°K. We will now consider what the composition, vertical depth and temperature distribution might be of the layer below the visible surface.

The studies by Lyot[13] and Kuiper[11] of the polarization of the reflected sunlight indicate that the visible surface of Venus is the top of a cloud of small particles suspended in the air. But it does not seem that this cloud of particles can be a water cloud.

In terrestrial water clouds, the change in the gradient of relative humidity that is observed in passing from the cloud to the air above is due, in most instances, to the change of the vertical gradient of temperature at the cloud top (a change to a more stable lapse rate above the cloud), and not to any large change in the vertical gradient of the water vapor density at the cloud top. Therefore, no large error would be made in estimating the amount of water

vapor immediately above the cloud if we assumed the vertical gradient of water vapor density to be constant, and assumed that within the cloud the air was saturated and the temperature lapse rate within the cloud was the saturated adiabatic rate. If we do this, and take the temperature of the cloud top as 237°K, then the lowest one kilometer layer of air above the cloud top will contain about 0·2 mm of precipitable water vapor. This is an order of magnitude larger than the water vapor observed on Venus.

But the matter does not rest there. For if, in fact, there is water vapor in the air above the cloud, then the cloud top temperature of 237°, computed for a pure CO_2 Venus atmosphere, would not be the equilibrium temperature. The additional atmospheric back radiation by the water vapor would result in a higher cloud top temperature, and with the higher temperature the equilibrium water vapor content of the air above the cloud would be larger than 0·2 mm. The increase would probably be by an order of magnitude or more, and the water vapor on Venus would then be relatively easily observed from the Earth.

In considering the depth of the visible cloud layer we must take notice, first, of the observations that have recently been made of radiation from Venus in the centimeter wavelengths (summarized in a review article by Mayer[14]). When interpreted as thermal radiation, these observations give an apparent black body temperature of the order of 600°K. But as centimeter radiation can penetrate a cloud that is opaque in the micron wavelengths, it has been suggested that the centimeter radiation comes from the solid surface of the planet, and that it is the solid surface which has the high temperature of 600°K.

It is clear that we cannot accept the centimeter radiation as being of thermal origin and at the same time believe that the clouds are water clouds. For this would imply a cloud layer that is incredibly thick. In an atmospheric layer that is either wholly or partly filled with water cloud the limiting lapse rate is approximately the adiabatic rate for saturated air, and this rate falls off rapidly with increasing temperature. To reach a temperature of 600°K, following the saturated adiabatic temperature curve downward from a cloud top temperature of 200° or 300°K, the cloud layer would have to be hundreds of kilometers thick. And a cloud layer hundreds of kilometers thick would not permit the downward penetration of the sunlight necessary to maintain the high surface temperature and the convection necessary to maintain the clouds.

If the clouds are not water clouds, the limiting lapse rate could be the adiabatic rate for a dry gas, and to reach the surface temperature of 600°K the cloud layer would not need to be more than 30–40 km deep. But even in this case it is difficult to see how the short wave absorption by the cloud would remain small enough for enough sunlight to penetrate the cloud to maintain the high surface temperature and the convection necessary to keep the cloud particles suspended. A greenhouse will not work, after all, if it is made of very dirty glass.

There is a further difficulty with a strong greenhouse effect—and this difficulty would exist even if, as has been suggested, the clouds are only a thin layer floating high in the atmosphere. Due to the spherical form of the planet, with a strong greenhouse effect there would be a correspondingly large horizontal gradient of heating from equator to pole. On a slowly rotating planet (though not to the same degree on a rapidly rotating one) a large horizontal heating gradient will result in a strong meridional circulation (the air rising at the equator and sinking at the poles) and the vertical heat transport by this large scale circulation will counteract the greenhouse effect even if the air is convectively stable. In this way not only small scale vertical convection but also the general circulation of the atmosphere places a limit on the greenhouse effect. It would require further study before one could say what this limit would be; but offhand, a greenhouse effect of 300° seems excessive.

Thus we see that there are serious theoretical difficulties in accepting the very high ground temperature that comes from interpreting the centimeter radiation as thermal radiation. We ought, therefore, to keep the possibility open that the centimeter radiation is of nonthermal

origin, and examine other evidence and clues as to what may lie below the visible surface.

The most striking feature of the radiometric scans of Venus, as indicated earlier, is the very small change from the day to the night side of the planet. In our model this implies a very small diurnal change in the temperature of the visible surface. We must note that there is nothing like this at the surface of either Mars or the land areas of the Earth.

Vaucouleurs[15], summarizing the radiometric measurements for Mars, gives a diurnal temperature change at the equator of about 55°C.

On the desert land areas of the Earth the diurnal variation of the ground temperature is also of the order of 50°C. It is sometimes stated that the large diurnal change of ground temperature in the Earth's deserts is due to the dryness of the air there, and hence due to a weak greenhouse effect. But this is not true. The amount of water vapor in the atmosphere over the Sahara is not less than it is over Europe. The diurnal surface temperature variation is large in the deserts because the *ground* is dry there and therefore a poor conductor of heat. And if the Earth would rotate more slowly the diurnal temperature change of the desert lands would be even larger than what is observed.

But over the oceans of the Earth the diurnal temperature change of the surface is very small; only some tenths of a degree. This is due to the large diurnal heat storage in the oceans.

The very simplest way to account for the observed normal decrease of temperature from equator to pole on Venus, but only a very small temperature change from day to night, is to assume that the visible cloud layer on Venus is very thin, of the character of a ground fog, and that beneath it there is a liquid layer, or ocean (but not of water). The diurnal heat storage in this ocean would maintain the temperature nearly constant day and night, even on a slowly rotating planet.

There is also other evidence which can be interpreted as indicating that the visible cloud is relatively thin and translucent, weakly revealing an ocean surface underlying it. Koval[16] and other Russian investigators have observed that the maximum of brightness in the infra-red, visible, and ultra-violet light corresponds to the position on Venus for which the angle of incidence and the angle of reflection of the Sun's rays are equal. These observations have been interpreted by Fesenkov[17] as indicating specular reflection by an ocean surface on Venus, blurred by an overlying thin cloud layer.

But if there is an ocean on Venus covered by a thin fog-like layer of cloud, and neither ocean nor cloud are made of water, what might their composition be? The only suggestion that the writer can make in answer to this question is that we seriously consider the hypothesis of F. Hoyle, that Venus is covered by an ocean of hydrocarbons.

Hoyle[18] believes that the matter out of which Venus accumulated contained water, just as did the matter out of which the Earth accumulated; but that on Venus, unlike the Earth, there was an excess of hydrocarbons over water. As a consequence, the oxygen, formed through photo-dissociation of the water vapor on Venus by ultra-violet light from the Sun, combined with the hydrocarbons and left no residual oxygen to shield the remaining water vapor and reduce the rate of dissociation, as on the Earth. All the water on Venus was therefore destroyed, the hydrogen escaping to space and the oxygen combining with the carbon compounds to form the carbon dioxide of the atmosphere of Venus. A brief outline of this possible evolution of a planetary atmosphere is also given by Urey[19].

Hoyle attributes the presumed slow rotation rate of Venus to the large tides raised in the hydrocarbon ocean by the gravitational action of the Sun, and the brilliant appearance of Venus (its high albedo) to a "smog" of droplets of oil covering the ocean surface.

The writer has only one comment to make on this hypothesis of Hoyle's, and that is that it is not necessary to assume that the amount of carbon compounds on Venus was initially greater than the amount of water, nor that all of the water has since disappeared from the planet. As low density hydrocarbons will float on water, he suggests that if there were simply enough hydrocarbons on Venus to cover the

water with a sufficiently thick layer, then the slow molecular diffusion of water upward through the hydrocarbons would be less than the rate at which the water in the atmosphere is dissociated by the Sun's ultra-violet light, and the atmosphere of Venus would be essentially water vapor free. The water will remain on the planet, but it will be trapped beneath the layer of hydrocarbons, with no effective way of getting into the atmosphere.

It may be such an ocean with an upper layer of hydrocarbons, and covered by a thin smog of oil droplets, that constitutes the visible surface of Venus. The smog covering must be thin enough to have only a small temperature difference between its bottom and top, so that its effective radiating temperature is essentially the same as that of the nearly constant day and night temperature of the underlying ocean.

5. THE GENERAL CIRCULATION OF THE VENUS ATMOSPHERE

The dynamics of the thermally driven circulation of a planetary atmosphere is so complex that we cannot deduce from first principles even the general form of the circulation without actually integrating the governing equations. The state of the art is such that we are only now learning how to perform the integrations which will tell us how an envelope of gas, gravitationally held to a rotating sphere and, say, initially at rest, will respond to a beam of parallel radiation from the Sun[20].

But pending such dynamical studies of what the circulation on Venus ought to be according to first principles, there may be some value in attempting to construct a rough empirical model of the circulation by inference from the observed patterns in the images of the planet as seen in visible and in ultra-violet light, aided by cautious analogy with observed conditions in our own atmosphere.

The appearance of Venus in visible light (the appearance of the visible surface) is usually described as almost, but not quite featureless. Some dark shadings are observed, which are faint and elusive. Moore[21] in summarizing the observations of Venus since the seventeenth century, says, "The most infuriating feature of the shadings is their impermanence. We cannot follow them as they are carried across the planet's disk by the axial rotation, as can be done in the case of Jupiter; they alter so quickly that they are difficult to recognize from one night to the next, and in any case they are hopelessly blurred and hazy."

But some order, nonetheless, has been made of these shadings by the French astronomers, notably Dollfus[22]. Dollfus superimposed the dark areas in his drawings of the planet for eight to ten successive nights, and found that the superimposed images formed coherent patterns, as shown, for five different phases of the planet, in Fig. 3(A). From these composite drawings he constructed the planisphere of the visible surface shown in Fig. 3(B), where the center of the projection is the sub-solar point of the planet.

What strikes us immediately is the cellular character of the pattern, quasi-symmetrical about the sub-solar point. It brings to mind the cellular character of the low-level circulation of the Earth's atmosphere, which is, however, symmetrical about the polar axis.

The cellular form of the low-level circulation in the Earth's atmosphere was first called attention to by Exner[23], who showed that on a rotating Earth the cellular form was required in order to provide the meridional wind components necessary for the poleward heat transport. In Exner's scheme, shown in Fig. 4(A), cyclones and anticyclones alternate with longitude. But later, Bjerknes[24] showed that in the sub-tropical latitudes there must be only anticyclones (or anticyclones predominating), as in Fig. 4(B), in order that the zonally averaged wind be easterly in the low latitudes and westerly in the middle latitudes. In Bjerknes' circulation scheme the shear lines at the boundaries of the anticyclones are lines of frictionally induced horizontal velocity convergence and ascending air, and in the central region of each anticyclone there is frictionally induced horizontal velocity divergence and descending air.

Either of these two circulation schemes can be used to interpret, by analogy, the planisphere

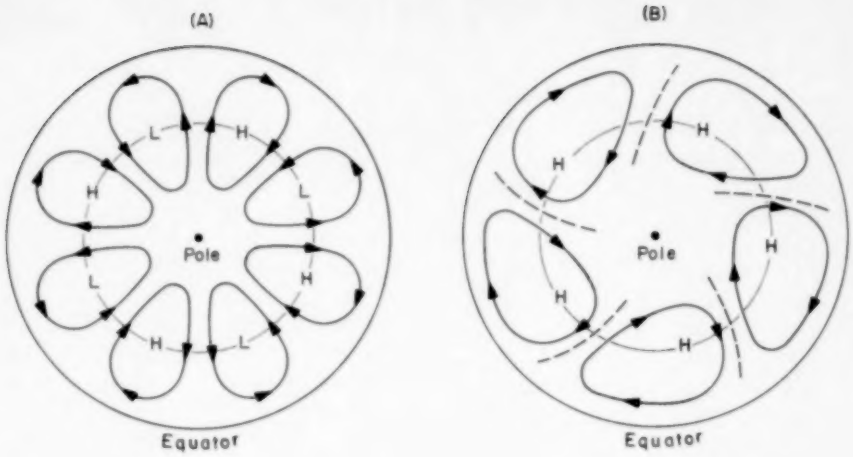


Fig. 4. Cellular form of the Earth's low-level atmospheric circulation: (A) after Exner (1925), (B) after Bjerknes (1935).

of Dollfus. In Fig. 5 we show an interpretation following the circulation model of Bjerknes. Assuming a direct sense of rotation of the planet, we have arranged a set of anticyclonic cells with the symmetry points for the cells not at the poles of rotation, but at the sub-solar and anti-solar points. These currents transport heat from low to high latitudes and, to a lesser extent, from the daytime to the night-time side of the slowly rotating planet. At the boundaries of the cells there is frictionally induced horizontal convergence and the air ascends. In the central regions of the cells there is horizontal divergence and the air descends.

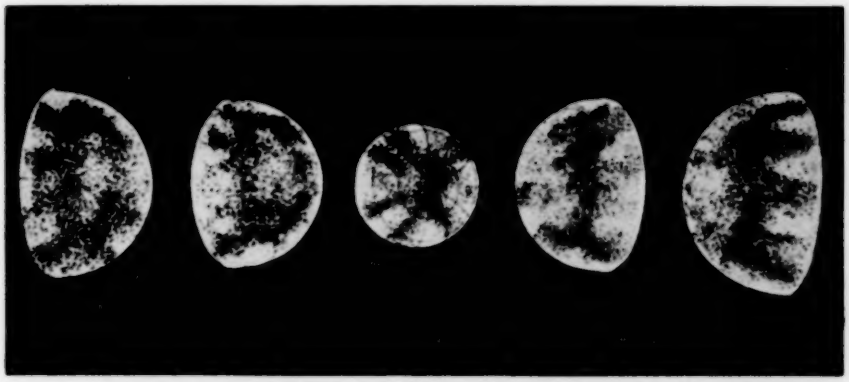
The sinking air is heated adiabatically and forms a weak "subsidence inversion" at the ocean surface. As shown in Fig. 5(C), the smog is trapped at the base of the inversion. But along the lines of convergence the lapse rate is less stable and the smog particles are diffused upward and dispersed by the higher level winds. Looking down on the planet in visible light, the observer would see both the brilliantly white smog and the darker smog-free (or less dense smog) areas in the cellular pattern shown in Figure 5(B). All this, of course, is pure inference by "indirect aerology," and with no attempt to determine whether such a circulation pattern, drifting with the sub-solar point over the slowly rotating planet, is dynamically possible.

By contrast, the photographs of Venus in ultra-violet light, by Ross[12] and others, show a structure of light and dark patches and bands more or less parallel to the equator of insolation of the planet. We have already interpreted this structure as due to the scattering of the Sun's ultra-violet light by the CO_2 cirrus clouds in the atmosphere high above the visible surface. From the zonal arrangement of the bands, we infer that the poleward temperature gradient, on the rotating planet, has produced in the higher levels of the Venus atmosphere a circumpolar vortex of the same form as that which we find at the higher levels over the Earth.

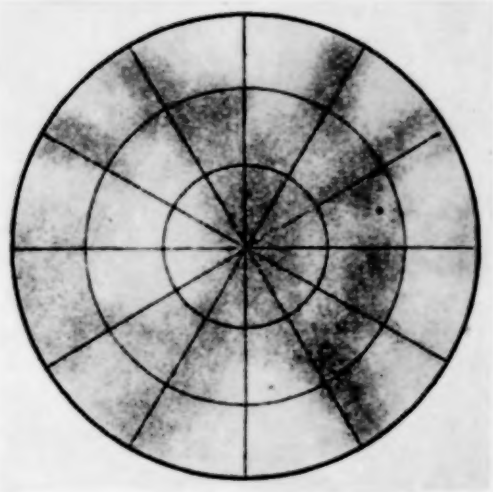
6. CONCLUSION

Although Venus is closer to the Sun than the Earth is, its larger albedo and the weaker greenhouse effect of its water-free atmosphere make its characteristic temperatures lower than on the Earth.

Unlike the Earth, where the radiative exchange by water vapor produces and supports a strong convective activity within the lower atmosphere, on Venus the radiative exchange



(A)



(B)

Fig. 3. (A) Composite observations of the visible surface of Venus, at different phases. (B) Planisphere of the visible surface (from Dollfus, 1955).

Refer to page 148
(facing page 150)

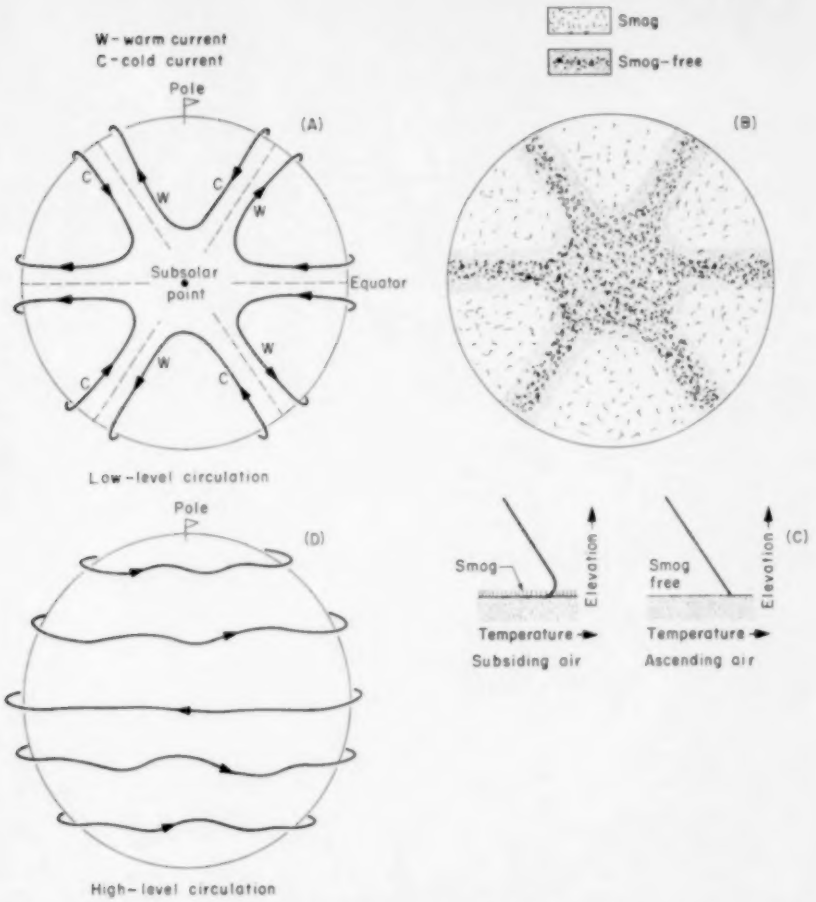


Fig. 5. Low-level cellular circulation and high-level zonal circulation of the Venus atmosphere, and distribution of the surface smog.

by carbon dioxide alone establishes a stably stratified atmosphere in which convection and turbulent air motions are suppressed.

The small difference of temperature between the day and night sides of Venus is attributed to a large heat capacity of the layer below the visible surface. From this we infer a universal ocean on Venus, but at least the uppermost part of this ocean is a liquid other than water.

From the markings seen in visible and in ultra-violet light, we infer a cellular form for the low-level atmospheric circulation, symmetrical about the sub-solar point, and a zonal form for the high-level circulation. The dark markings seen in visible light are where convergence lines in the low-level circulation disturb the thin smog that covers the ocean; it being the smog that gives the planet its high albedo. The bright markings seen in ultra-violet light are cirrus clouds formed by sublimation of carbon dioxide in ascending currents at high levels in the atmosphere.

Acknowledgements—The writer thanks the many persons with whom he had stimulating and useful discussions while preparing this paper. In particular, he thanks Dr. William W. Kellogg for his interest and assistance.

REFERENCES

1. G. Herzberg, *Atmospheres of the Earth and Planets.* edited by G. P. Kuiper, pp. 406-416. Revised Edition, Univ. Chicago Press, Chicago (1952).
2. G. S. Callender, *Quart. J. R. Met. Soc.*, **67,** 263 (1941).
3. G. Ohring *J. Met.*, **15,** 440 (1958).
4. J. N. Howard, D. L. Burch and D. Williams, *Geophysical Research Papers*, No. 40, Geophys. Res. Dir., AFCRC Cambridge, Mass. (1955).
5. N. A. Kozyrev, *Astrophys. Obsns. Crimea,* **12,** 177 (1954).
6. H. N. Russell, *Astrophys. J.*, **43,** 173 (1916).
7. W. M. Sinton, *Distribution of Temperatures and Spectra of Venus and other Planets.* Doctoral Dissertation, Johns Hopkins University, (1953).
8. W. M. Sinton and J. Strong, *Astrophys. J.*, **131,** 470 (1960).
9. E. Pettit and S. B. Nicholson, *Publ. Astr. Soc. Pacific*, **67,** 293 (1955).
10. J. W. Chamberlain and G. P. Kuiper, *Astrophys. J.*, **124,** 399 (1956).
11. G. P. Kuiper, *The Threshold of Space*, edited by M. Zelikoff, pp. 78-86 Pergamon Press, New York, (1957).
12. F. E. Ross, *Astrophys. J.*, **68,** 57 (1928).
13. B. Lyot, *Recherches sur la Polarisation de la Lumière des Planetes et de Quelques Substances. Terrestres. Ann. Obs. Paris (Meudon)*, **8,** Paris (1929)
14. C. H. Mayer, *Astron. J.*, **64,** 43 (1959).
15. G. de Vaucouleurs, *Physics of the Planet Mars.* Faber and Faber, London, (1954).
16. I. K. Koval, *Astron. Journ. Acad. Sciences USSR*, **2,** No. 5, 792 (1958).
17. A. Oparin and V. Fesenkov, *The Universe*, Foreign Languages Publishing House, Moscow, (1957).
18. F. Hoyle, *Frontiers of Astronomy.* Harper and Bros., New York (1955).
19. H. Urey, *Science in Space,* Nat. Acad. Sci., Washington (1960).
20. *The Dynamics of Climate,* (*Proceedings of the General Circulation Conference, Institute for Advanced Study at Princeton;* Edited by R. Pfeffer), Pergamon Press, London and New York, (1960).
21. P. Moore, *The Planet Venus*, Macmillan Co., New York, (1957).
22. A. Dollfus, *L'Astronomie*, **69,** 413 (1955).
23. F. M. Exner, *Dynamische Meteorologie*, pp. 215-218, 2nd ed., Verlag Julius Springer, Vienna, (1925).
24. J. Bjerknes, *Scientia*, **57,** 114 (1935).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 153–157. Printed in Great Britain.

TEMPERATURE MEASUREMENTS FROM OH BANDS

G. KVIFTE

Department of Physics, Agricultural College of Norway, Vollebekk, Norway

(*Received* 29 *July* 1960)

Abstract—A survey of temperature measurements from OH rotation vibration bands (Meinel system) by different observers is given. Methods of measurement are discussed.

It is pointed out that the temperatures given by some observers have to be corrected because an incorrect rotational constant has been used.

New data obtained at Ås, Norway, confirm earlier observations and yield a temperature of the atmospheric OH-emitting layer of about 220°K at this geographical latitude (59·7°N).

The different observers' corrected temperatures are discussed with a view to ascertaining if the alleged latitude effect is real. It is concluded that the OH temperature is fairly constant at geographical latitudes between 40 and 60°N with a mean value of about 220°K and seems to rise northwards from say 60°, reaching a temperature of approximately 300°K at 75°N.*

1. INTRODUCTION

During the last ten years, since Meinel[1] published the discovery of the rotation–vibrational bands of the OH molecule in the nightglow spectrum, several observers have studied these bands with a view to determining the temperature of the OH-light emitting layer of the atmosphere. The rotational lines in the *P*-branches of the bands are fairly well spaced and can be easily obtained separated with a nightglow spectrograph of medium dispersion. The experimental circumstances should therefore be favourable for accurate temperature measurements. The published results are, however, rather divergent, as will be seen in Table 1 where experimental data and the computed temperatures of the different investigators are assembled.

The results vary from 150°K to 350°K. Some of this variation is probably due to inadequacy of the spectral instruments, some to the method which was used in evaluating the temperature. The temperatures are, in recognition of these facts, only given as rough estimates by some authors.

Those values, which are believed to be more accurate, also differ appreciably. Opinions have been advanced that the OH rotational temperature may vary with geographical latitude[2] or season[3] and that the divergent results may be explained in this manner. It is felt that a critical discussion of the precision of the measurements is needed in order to try to ascertain whether the differences are real or not.

2. INSTRUMENTAL REQUIREMENTS

The OH rotation–vibrational band system is a doublet system, $^2\Pi$–$^2\Pi$. It is only the *P* branches which can be used with any success in temperature evaluation. The spacing between corresponding lines in sub-branch $P_1(^2\Pi_{3/2}$–$^2\Pi_{3/2})$ and sub-branch $P_2(^2\Pi_{1/2}$–$^2\Pi_{1/2})$ varies from say 5 to 10 Å for actual quantum numbers in the red and the near infra-red region. A precise measurement of the rotational temperature is only possible when these lines are well separated.

When the spectrum is photographed on high speed emulsions, usually necessary in order to shorten the exposure times, the grains in the emulsion set the limit of the resolving power to about 50 lines/mm. Combined with the line spacings mentioned above, the spectral instrument must have a dispersion corresponding to a linear scale value of not more than 250 Å/mm in the red region and not above

* *Note added in proof*: A paper by Dr. Wallace (*J. Geophys. Res.* **65**, 3, 1960) dealing with correction of OH temperatures, has recently been brought to the author's attention. Dr. Wallace's conclusions agree well with those in this paper.

Table 1. Temperatures, T, Derived from OH Bands

Author	Lat. (°N)	Instrument	$(d\lambda/ds)$ Å/mm (Slit width Å)	Region*	T °K	Method
Meinel	42·6	Sp. gr.		I.R.	260 ± 5	$I - J$ curve
Cabannes *et al.*	43·9	,,	350	R	150 – 250	,,
Dufay *et al.*	43·9	,,	700–800	I.R.	255 – 270	,,
Chamberlain *et al.*	76·6	,,	140– 70	R	300 – 350	,,
Gush *et al.*	52·1	Photom.	(100 Å)	F.I.R.	200 ± 50	,,
Mironov *et al.*	51·05	Sp. gr.	80–130	R., I.R.	230 ± 20	Not stated
Mironov *et al.*	63·6	,,	80	R	280 – 330	,,
Dufay M.	43·9	Photom.	(15 Å)	I.R.	250	$I - J$ curve
Kvifte	59·7	Sp. gr.	35	R	215 ± 5	In $I - E$ curve
McPherson *et al.*	52·1	,,	120– 60	I.R.	234 ± 29	,,
McPherson *et al.*	58·8	,,	60		246	
McPherson *et al.*	74·7	,,	120	,,	296 ± 38	,,

* R = red, I.R. = near infra-red, F.I.R. = far infra-red.

500 Å/mm in the near infra-red region if the sub-branches are to be barely separated. These are conservative values, and the dispersion ought to be doubled at least to give reliable intensity measurements of the lines.

The continuous light present in the airglow gives a more intense background on the plates the lower the dispersion. Good intensity measurements of lines from spectral plates are contingent upon a low background intensity. It is the experience of the author that the scale value of the spectral apparatus should not be much greater than 100 Å/mm if the intensity measurements of near lying lines are not to be seriously impaired. The disturbance caused by overlapping of emission or absorption lines nearby must also be taken into account here.

One instrumental requirement is then that the scale value of the spectrograph should be better than, say, 100 to 150 Å/mm. Another is that the resolving power should be sufficient to take full benefit of the resolving power of the plate. This condition is usually fulfilled by spectrographs.

The resolving power of photometers is, however, most often limited by the exit slit to the photocell. When used in temperature measurements, the exit slit should be less than, say, 5 Å. Otherwise special evaluation techniques should be used.

When eliminating those temperatures in Table 1 which are obtained with instruments not fulfilling the requirements here stated, the extreme values disappear.

3. EVALUATION METHODS

The intensity, I, of a rotational line of a band from a radiating gas in thermal equilibrium is given by:

$$I(J) = Cs_J \nu_J^4 \exp(-F(J)\, hc/kT) \qquad (3.1)$$

ν_J is the wave number of the line, s_J the line strength, J the rotational quantum number of the upper vibrational level and $F(J)$ the rotational term value of the same level. T is the equilibrium temperature, and the constants h, c and k have the usual meanings.

Theoretical expressions for s_J are given by Hönl and London[4]. $F(J)$ is often approximated by

$$F(J) = B_{et} J(J+1) \qquad (3.2)$$

where B_{et} is the "effective" rotational constant of the upper vibrational level.

One method for determining the temperature T is to plot the measured intensities I of different rotational lines against the rotational quantum number J. Theoretical intensities, computed from equation (3.1) for different values of the parameter T, are plotted on the same graph, the constant C being adjusted to make experimental and theoretical intensities coincide for a specified quantum number. The

temperature is then fixed as the parameter to the theoretical curve which has the best fit with the measured intensity curve.

This method, called I–J/curve method in Table 1, is laborious and not very accurate.

A more direct procedure, and presumably a more precise one, is the ln (I)–E method. Here $\ln(I/sv^4)$ when plotted against the energy $E(J)=F(J)\cdot hc$ should give a straight line if thermal equilibrium exists. The temperature is computed from the slope of the straight line.

When $F(J)$ is approximated by the expression (3.2), the plotting is done against $J(J+1)$. Designating the gradient of the $\ln(I/sv^4)-J(J+1)$ curve by $(-\alpha)$, then T is given by

$$T=\frac{B_{ef}\cdot hc}{k\cdot\alpha} \tag{3.3}$$

It is to be noted here that B_{ef} is different for the $^2\Pi_{3/2}$ and the $^2\Pi_{1/2}$ states of the OH molecule, a fact which seems to have been generally overlooked in temperature computations. For the rotational constants given in the literature (e.g.[5]) B_v corresponds to the unsplit $^2\Pi$ state. The effective rotational constants B_1 and B_2, belonging to the $^2\Pi_{3/2}$ and $^2\Pi_{1/2}$ states respectively, may be found from

$$B_{ef}=B_v\left(1\pm\frac{1}{Y}\right) \tag{3.4}$$

where Y is a measure of the coupling between the spin and the orbital angular momentum[6], and + and − correspond to B_1 and B_2 respectively.

However, equation (3.4) only gives approximate values. When $F(J)$ is known for the two sublevels, a more correct set of the rotational constants may be computed from equation (3.2). Table 2 contains values of B_v, of B_{ef} computed from equation (3.4) and of B computed from equation (3.2) as a mean from five rotational levels.

It is seen that B_v differs by about 10 per cent from B_{ef} and about 8 per cent from B. When using B_v instead of B_{ef} or B in equation (3.3), the resulting temperature will have a per cent error of the same magnitude.

The mean of B_1 and B_2 is nearly equal to B_v. It might be argued that if the mean of the intensities of corresponding lines in the P_1 and P_2 branches were used, B_v should be the correct value to insert in equation (3.3). This is not so because the two branches are far from equal in intensity, and therefore enter in the mean with different weight.

Table 2. Rotational Constants of Vibrational Levels in the $^2\Pi$ Ground State of OH

	B_v*	$B_{ef}=B_v(1\pm1/Y)$		B† $=F(J)/J\cdot(J+1)$	
v	$^2\pi$	$^2\pi_{3/2}$	$^2\pi_{1/2}$	$^2\pi_{3/2}$	$^2\pi_{1/2}$
0	18·515	16·05	20·95	16·93	20·19
1	17·807	15·55	20·10	16·30	19·41
2	17·108	15·04	19·20	15·70	18·69
3	16·414	14·5	18·36	15·07	17·93
4	15·722	14·0	17·50	14·49	16·98
5	15·033	13·41	16·65	13·88	16·20
6	14·348	12·89	15·80	13·23	15·31
7	13·673	12·33	15·00	12·68	14·47
8	12·999	11·78	14·20	11·99	13·69
9	12·328	11·23	13·43	11·31	12·80

* From Herman and Hornbeck[5]
† $F(J)$ from Kvifte[7]

Although B_{ef} or B used in connection with equation (3.3) should give acceptable temperature values, a better result is obtained by plotting $\ln(I/sv^4)$ directly against the experimental values of $F(J)$ (or rather $F(J)\cdot hc/k$). Approximations are then avoided. The author has used this procedure enclusively in his temperature measurements.

4. TEMPERATURE MEASUREMENTS FROM NORWAY

During the winters of 1957–1960 a number of nightglow spectra were obtained at Ås, Norway, with a grating spectrograph. Several of these were suitable for temperature measurement. As instrumental data and measuring technique have been dealt with elsewhere[7] only some of the results will be given and discussed here.

Table 3, giving temperatures derived from two spectra with four measurable bands each, illustrates the relatively good correspondence between the temperatures computed from the

Table 3. Temperatures (°K) measured from OH Bands, Norway 1959

		Upper*		Lower*		U+L		
Date	Band	P_1	P_2	P_1	P_2	P_1	P_2	Mean
3.2.59	5·1	225	221	221	221	223	223	222
	6·2	219	218	219	224	219	219	219
	8·3	215	215	216	217	216	216	216
	9·4	236	217	240	217	240	217	228
	Mean	224	218	224	220	224	219	221
5.2.59	5·1	217	222	220	222	218	222	220
	6·2	217	215	217	215	215	211	215
	8·3	237	217	230	215	231	214	224
	9·4	223	227	221	218	215	221	221
	Mean	224	220	222	218	220	217	220

* From microphotometer tracing along the upper and the lower part respectively of the spectrum.

two sub-branches P_1 and P_2. The P_1-temperatures are on an average 4°K higher than the P_2-temperatures. It should be noted that if the rotational constant B_v had been used in equation (3.3) in evaluating the temperature from both branches, the difference would have been about 35°K more, giving a total difference of about 40°K between the P_1 and P_2 temperatures.

In Table 4 all temperatures measured at Ås, Norway, are collected. The values are fairly consistent and give an average temperature of 220°K for two winters. It should be noted that the apparent errors in the temperatures, estimated by the standard deviation, s, in the last column of the table, increase rapidly with an increasing scale value for the spectrograph. This observation supports the discussion in Section 2.

5. CONCLUSIONS

As mentioned in Section 3 some observers have not been aware of the necessity of using different rotational constants when evaluating temperatures from the two sub-branches in the OH Meinel system. In those cases where the intensities of the rotational lines have been given, the temperatures have been recomputed by the procedure outlined in Section 3.

Table 4. Temperatures (°K) Measured from OH Bands, Norway 1958–1959

Date	Band	$d\lambda/ds$ Å/mm	T(°K)	T_m(°K)	Mean $\pm s$*
16.1.58	6·1	35	214	215·5	215 ± 1·8
,,	9·3		217		
17.1.58	6·1		215	215	
20.1.58	6·1		215	215	
21.1.58	6·1		214	214	
13.2.58	6·1		213	214	
,,	9·3		215		
17.2.58	6·1		215	216·5	
,,	9·3		218		
13.3.58	6·1		215	215	
27.3.58	6·1		214	212·5	
,,	9·3		211		
3.2.59	5·1	80	222	221	220 ± 4·5
,,	6·2		219		
,,	8·3		216		
,,	9·4		228		
5.2.59	5·1		220	220	
,,	6·2		215		
,,	8·3		224		
,,	9·4		221		
21.2.59	6·1	100	216	220	224 ± 15
,,	9·3		224		
8.3.59	6·1		227	240	
,,	9·3		252		
10.3.59	6·1		220	228	
,,	9·3		235		
12.3.59	6·1		207	208	
,,	9·3		209		

* Standard deviation.

Table 5 contains the temperatures, T, given by different authors and the corrected values, T_{corr}. A sign of interrogation after the values in this column indicates that data were not available to help decide as to whether the temperature should be corrected or not. Temperatures obtained from instruments not fulfilling the requirements stated in Section 2 have been put in round brackets. The entries have been placed in order of increasing geographical latitude for the stations of observation.

In the group of observations from latitudes around 43°N, Meinel's values are probably the most precise. He himself gives an error of ±5°K. The corrected temperature, 241°K, computed as a mean of the temperatures derived from his 14 single observations has,

Table 5. Latitude Variation of OH Temperatures

Author	Number of Measurements	T (°K)	$T_{corr.}$ (°K)	T_m	Lat. (°N)
Meinel	14	260	241	240?	42·6
Cabannes *et al.*		150–250	(200)		43·9
Dufay J. *et al.*	6	255–270	(240)		43·9
Dufay M.	2	250	240		43·9
Mironov *et al.*		230±20	230?	220	51·1
Gush *et al.*		200± 50	(200)		52·1
Mc Pherson *et al.*	15	234	217		52·1
Mc Pherson *et al.*	2	246	230	220	58·8
Kvifte	12	215	215		59·7
Kvifte	16	222	222		59·7
Mironov *et al.*	2	280 Z, 330 N	280?	280?	63·6
Mc Pherson *et al.*	4	296	273	300?	74·7
Chamberlain *et al.*	4	300–350	325?		76·6

however, a standard deviation of about 20°K. Therefore, although three of the four means in this group coincide, the actual temperature may well differ appreciably from it.

The next two groups, around latitudes of 52°N and 59°N, have the same mean, viz. 220°K. The temperatures given by the different authors are based on a number of observations and the agreement between them is satisfactory. The temperature means seem to be reliable.

Mironov *et al.* have measured a temperature of 280°K in the zenith direction at 63·6°N. The method employed has not been given nor the intensities stated. If a correction were applied to the values, the temperature would have to be lowered to about 260°K.

The two temperature estimates around latitude 75°N seem somewhat uncertain. Chamberlain *et al.* indicate a temperature between 300 and 350°K. These temperature limits have not been corrected owing to lack of data. McPherson *et al.* state a standard deviation of about 40°K from their mean value, 296°K. The corrected value, 273°K, however, has a standard deviation of 25°K. The average temperature, 300°K, of this latitude may well be wrong by, say, 25 to 30°K.

These considerations seem to justify the following conclusion: The temperature of the OH-layer in the atmosphere is, globally viewed, fairly constant at medium geographical latitudes, say between 40 and 60°N, with a mean at about 220°K. There is an indication that the temperature rises for higher latitudes by some 75°K from about 60°N to 75°N.

It is felt, however, that more precise data are needed to verify this effect. Simultaneous observations at different latitudes with nearly identical instruments would be preferable. Such observations are intended to be carried out in Norway next winter.

REFERENCES

1. A. B. Meinel, *Astrophys. J.* **111,** 555 (1950).
2. A. V. Mironov, V. S. Prokudina and N. N. Shefov, *Ann. Géophys.* **14,** 364 (1958).
3. D. H. McPherson and A. Vallance-Jones, *J. Atm. Ter. Phys.* **17,** 302 (1960).
4. H. Hönl and F. London, *Z. Phys.* **33,** 803 (1925).
5. R. C. Herman and G. A. Hornbeck, *Astrophys. J.* **118,** 214 (1953).
6. G. Herzberg, *Spectra of Diatomic Molecules.* Van Nostrand Co., N.Y. (1951).
7. G. Kvifte, *Geofys. Publ.* **20,** No. 12 (1959).
8. J. Cabannes, J. Dufay and M. Dufay, *C.R. Acad. Sci. Paris,* **230,** 1233 (1950).
9. J. W. Chamberlain and N. J. Oliver, *Phys. Rev.* **90,** 1118 (1953).
10. J. Dufay and M. Dufay, *C.R. Acad. Sci. Paris,* **232,** 426 (1951).
11. M. Dufay, *Ann. Géophys.* **15,** 134 (1959).
12. H. P. Gush and A. Jones-Vallance, *J. Atm. Ter. Phys.* **7,** 285 (1953).
13. G. Kvifte, *J. Atm. Ter. Phys.* **16,** 252 (1959).
14. A. B. Meinel, *Astrophys. J.* **112,** 120 (1950).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 158–162. Printed in Great Britain.

RELATIVE POPULATION OF OH X $^2\Pi$ LEVELS

G. KVIFTE

Department of Physics, Agricultural College of Norway, Vollebekk, Norway

(*Received* 29 *July* 1960)

Abstract—Intensities of nightglow bands in the visible and the near infra-red region of the OH rotation vibration system are presented. Comparison with other similar measurements shows a fair consistency, and the absolute intensities given do not, except in a couple of cases, deviate seriously from those predicted previously.

The relative population of the vibrational levels $v=3$ to 9 of the OH ground state is computed from the intensities by means of relative transition probabilities given by some previous investigators. The values obtained are, in descending order, from $v=9$: 1·0, 1·25, 1·9, 2·6, 5·1, 7·0 and (8·1). The transition probabilities given by one investigator give somewhat larger values.

Relative excitation numbers of the different levels have been evaluated from the relative population. It is concluded that excitation of the vibrational levels of the OH ground state takes place at all levels and that the relative number of molecules excited at different levels is about the same.

The bearing of these results on the theories of Bates, Nicolet and Herzberg and of Krasovskii is discussed.*

1. INTRODUCTION

Any theory of the excitation of the OH rotation-vibration system in the nightglow has to account for the fact, discovered by Meinel[1], that emission from levels higher than $v_1=9$ is absent or very weak. Two mechanisms have been proposed, both of which seem to explain this interesting observation.

Bates and Nicolet[2] and Herzberg[3] have independently suggested the following cycle:

$$H+O_3 \longrightarrow OH^* + O_2\ (+3{\cdot}30\text{eV}) \qquad (1.1)$$

$$OH+O \longrightarrow O_2+H \qquad (1.2)$$

The energy released in process (1.1) is just sufficient to excite the ninth vibrational level of OH. If the OH emission takes place at a height of 70–80 km in the atmosphere, the process seems to furnish a sufficient number of excited OH molecules to account for the intensity of the OH emission.

Krasovskii[4] has advanced an alternative hypothesis. He considers the process

$$H+O_2^* \longrightarrow OH^* + O\ (-0{\cdot}75\text{eV}) \qquad (1.3)$$

in connection with the phenomenon (1.2) more likely than the mechanism (1.1)–(1.2). The oxygen molecule in the process (1.3) must be excited up to, but not above, the 27th vibrational level of the ground state ($x^3\Sigma$) if the 9 lower levels of the OH molecule, and no more, should be activated. Krasovskii's process should take place at the 100 km level in the atmosphere.

Arguments, *pro et contra*, have been forwarded regarding both hypotheses, but no decision has been reached. It is possible that knowledge of the relative population of the OH vibrational levels might be of some value for deciding between the theories. An evaluation of the relative population has been tried here.

2. INTENSITIES OF THE OH BANDS

Meinel[1] and M. Dufay[5] have measured relative intensities of several OH bands in the infra-red. Meinel's list of band intensities also includes measurements by Cabannes, J. Dufay

* When preparing this paper the author was not aware of the important paper by J. W. Chamberlain and C. A. Smith on the excitation rates and intensities of OH in the airglow (*J. Geophys. Res.* **64**, 611, 1959). It is very satisfying that the results given here are very similar to those obtained by these authors from other experimental data and slightly different methods.

and M. Dufay[6]. Fedorova[7] and Kvifte[8, 9] have given estimates of absolute intensities of bands in the infra-red and red regions. The procedure of the last named author for evaluating the intensities is briefly outlined below.

The relative intensity, Q_v, of a band (transition between vibrational levels v_1 and v_2) is the sum of the relative intensities, $I_v(J)$, of all rotational lines in all branches. The relative intensities of only a limited number, n, of the rotational lines is, however, measurable.

With

$$I_t(J) = Csv^4 \exp(-F(J)\, hc/kT) \qquad (2.1)$$

and

$$Q_t = \sum_J I_t(J), \qquad (2.2)$$

the sum being taken over all J (and all branches), Q_e is determined by:

$$Q_e = Q_t \cdot \sum^n I_e(J) / \sum^n I_t(J) \qquad (2.3)$$

The sums in equation (2.3) are taken over the n measured intensities, $I_e(J)$, and the corresponding theoretical ones, $I_t(J)$. All constants in the expression (2.1) are available. The temperature, T, has been put equal to 215°K.

The relative intensities Q_e have been standardized against the absolute intensity of the Na I-lines and to some extent against the intensity of the red OI doublet. The absolute sum intensity of the Na I and OI doublets have been taken from Roach[10].

The intensities obtained in this way in Norway are contained in the last row of the

Table 1. Intensities (R) of OH Bands

(v_1/v_2)	0	1	2	3	4	5	Author
							Meinel
3	720						Fedorova
	3700						Dufay M
							Kvifte
	300						M.
4		6400					F.
	340	9600					D
	330						K.
	30	1500					M.
			11200				F.
5		1150					D.
	35	1100					K.
		90	2800				M.
6			1200				F.
			1700				D.
		120	1100				K.
			190	7500			M.
7				1800			F.
				3000			D.
			280				K.
			90	450			M.
8					2300		F.
				500	2100		D.
			80	550			K.
				150	850		M.
9						2600	F.
					670	4550	D.
				160	900		K.

v_1-rows of Table 1. The table also contains values from Meinel, Fedorova and M. Dufay. The relative intensities given by Meinel and Dufay have been standardized by means of the Norwegian results. Fedorova's absolute intensities seem to be too low by a factor four, and have been corrected accordingly.

The intensities in the table give a fairly consistent picture and do not deviate seriously, except in a couple of cases, from the predicted intensities given by Roach[11].

3. RELATIVE POPULATION OF VIBRATIONAL LEVELS

The intensity of an emission line, I_{nm}, corresponding to a transition from state n to m, with wave number ν_{nm}, is given by

$$I_{nm} = N_n hc \nu_{nm} A_{nm} \tag{3.1}$$

N_n is the number of molecules in the initial state, and A_{nm} the transition probability. Expression (3.1) is also valid for an emission band if A_{nm} signifies the band transition probability and ν_{nm} is an average wave number of the band.

The population of vibrational levels of a molecular state, relative to the population of a specified level, may now be calculated by means of equation (3.1) if the band intensities and the relative transition probabilities are known.

It is appropriate for the OH molecule to compute the population relative to the 9th vibrational level according to the expression

$$\frac{N_n}{N_9} = \frac{I_{nm}}{I_{9f}} \cdot \frac{\nu_{9f}}{\nu_{nm}} \cdot \frac{A_{9f}}{A_{nm}} \tag{3.2}$$

Relative transition probabilities for the OH rotation-vibrational system have been computed by Shklovskii[12] and Heaps and Herzberg[13] independently. Their values do not differ significantly.

Table 2 contains the relative population of the vibrational levels $v=3$ to 9 computed according to equation (3.2) from the intensities in Table 1. Both sets of transition probabilities have been used. In view of the difficulty of measuring intensities exactly the values are satisfactorily concordant. Shklovskii's transition probabilities give altogether higher population values than those of Heaps and Herzberg. The last two columns in the table contain average values.

4. DISCUSSION

Heaps and Herzberg have computed the relative population of the vibrational levels of the OH ground state on the assumption that all OH molecules are excited to the 9th level and lose energy only by successive radiative transitions to the lower levels. Their values are given in the third column ($v_1=9$) of Table 3. The correspondence is rather unsatisfactory between these values and the population numbers computed from band intensities, entered in the second column of the table. If it is assumed that all levels are excited to the same extent and that depopulation only occurs as radiative transition to lower levels, much better correspondence can be obtained between experimental and theoretical values as seen by comparing the entries of the second and fourth

Table 2. OH T X $^2\Pi$. Relative Population Computed from Band Intensities

v	Meinel		Fedorova		Dufay		Kvifte		Mean	
	A_s*	A_H†	A_s	A_H	A_s	A_H	A_s	A_H	A_s	A_H
9	1	1	1	1	1	1	1	1	1	1
8	1·45	1·35	1·6	1·35	1·25	1·07	1·4	1·28	1·45	1·25
7	(3·8)	(3·6)	1·8	1·8	2·6	2·4	1·6	1·5	2·0	1·9
6	3·5	3·1	2·7	2·5	3·2	2·8	2·1	1·9	2·9	2·6
5	5·5	4·8	6·7	6·5	6·0	5·0	4·3	3·9	5·6	5·1
4	6·9	5·9	8·0	8·0	8·9	8·0	7·1	6·1	7·7	7·0
3			(3·6)	(3·1)	15·5	13			(9·5)	(8·1)

* A_{nm} from Shklovskii.
† „ „ Heaps and Herzberg.

Table 3. Relative Population and Excitation of OH $^2\Pi$ Vibrational Levels

v	From Band Intensities	Excitation to:			
		$v_1=9$	All*	All†	Weight‡
9	1·0	1·0	1·0	1·0	1
8	1·25	0·55	1·55	1·3	0·75
7	1·9	0·66	2·28	1·9	0·75
6	2·6	0·76	3·21	2·6	0·75
5	5·1	0·89	4·43	4·3	1·5
4	7·0	1·09	6·11	6·3	1·5
3	(8·1)	1·43	8·73	9·3	1·5

* Equal weight.
† Weighted.
‡ Weight = relative excitation number.

columns of Table 3. In computing the numbers in column 4 the transition probabilities of Heaps and Herzberg have been used, and the excitation processes (1.1) or (1.3) are supposed to furnish the same relative number of molecules excited in all vibrational levels up to $v=9$. The computation procedure is analogous to that used by Heaps and Herzberg for excitation to one level only.

It is, however, also possible to evaluate the relative excitation numbers of the different levels, i.e. the number of molecules which are excited directly by the processes (1.1) or (1.3) in each level. The relative excitation numbers in the last column of Table 3 give a relative population, shown in the last column but one, which agrees well with the experimental population (column 2).

Undue confidence should not be placed in the relative excitation numbers. There may be a considerable percentage error. It seems, however, to be a fairly probable conclusion that excitation of the vibrational levels of the OH ground state takes place to all levels and that the relative abundance of molecules excited in the different levels is about the same. There is an indication of a somewhat smaller excitation in the levels $v=6$ to 8 than in level $v=9$ and the lower levels. But this effect needs confirmation.

The conclusion seems at first to lend support to Krasovkii's hypothesis. The oxygen molecules active in the process (1.3) are supposed to be excited to vibrational levels up to $v=27$. This is sufficient to excite the OH molecule to the 9th vibrational level. The lower OH levels should then be activated by oxygen molecules excited to levels lower than the 27th, and presumably to about the same degree. It should, however, be noted that whereas the level OH ($v=9$) might be activated through the O_2 levels $v=27$ and 26, the OH ($v=8$) level has three possibilities, viz. O_2 $v=25$, 24 and 23, OH ($v=6$) two again etc. One should therefore believe that the OH ($v=8$) level should have a larger relative excitation number than the 9th level, the more so if, as Krasovskii[14] has stated, the relative abundance of oxygen molecules increases with decreasing vibrational excitation. According to the entries in Table 3 the relative excitation number of the 8th vibrational level is smaller than that of the 9th level.

Too little, however, is known at the moment about the rate coefficients of the process (1.3) for different excitation levels of the oxygen molecule; neither is the relative abundance of the differently excited molecules known. The objection may therefore be of little importance.

The Bates–Nicolet–Herzberg process is sometimes believed to be a resonance process exclusively exciting the OH molecule to the 9th vibrational level. If so, the data presented here contradict the theory. It is, however, the author's opinion that this might not be so. The possibility should exist, that part of the energy released in reaction (1.1) might be taken up by the oxygen molecule as excitation energy to a low vibrational level. The remaining energy then provides the excitation of the OH molecules to levels lower than the 9th; thus:

$$O_3 + H \rightarrow OH\,(v \leqslant 9) + O_2\,(v \geqslant O) \quad (4.1)$$

Such a reaction might well account for the observed OH band intensities.

REFERENCES

1. A. B. Meinel, *Astrophys. J.* **112,** 120 (1950).
2. D. R. Bates and M. Nicolet, *J. Geophys. Res.* **55,** 301 (1950).
3. G. Herzberg, *J. Roy. Astron Soc. Canada* **45,** 100 (1951).
4. V. I. Krasovskii, *Prog. Phys. Sci.*, **47,** N4, 493 (1952).

5. M. Dufay, *Ann. Géophys.* **15,** 134 (1959).
6. J. Cabannes, J. Dufay, and M. Dufay, *C. R. Acad. Sci.* **230,** 1233 (1950).
7. N. I. Fedorova, *Ann. Géophys.* **14,** 365 (1958).
8. G. Kvifte, *Geofys. Publ.* **20,** No. 12 (1959).
9. G. Kvifte, *J. Atm. Ter. Phys.* **16,** 252 (1959).
10. F. E. Roach, *I.A.G.A. Bul.* No. 15 b, 100 (1956).
11. F. E. Roach, *An. I.G.Y.* **4,** 115 (1957).
12. I. S. Shklovskii *An. I.G.Y.* **4** (1957).
13. H. S. Heaps and G. Herzberg, *Z. Phys.* **133,** 48 (1952).
14. V. I. Krasovskii, *The Airglow and the Aurorae,* 193. Ed. by Armstrong and Dalgarno (1956).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 163–167. Printed in Great Britain.

RESEARCH NOTES

Airglow of Venus: a re-examination

A comparison has been made of spectra of the bright crescent and of the unilluminated portion of Venus in the region of 4300 A to attempt to obtain the spectrum of the night sky of Venus during a four-month interval including the 1959 inferior conjunction in order to confirm the previous results obtained by N. Kozyrev[1] and G. Newkirk[2] and to attempt to determine the nature of the reported emission features.

We modified the instrument used by Newkirk in an earlier investigation of the airglow of Venus[2] by increasing the dispersion of the spectrograph and by artificially widening the spectra by means of a rocking, plane-parallel plate of glass immediately behind the spectrograph slit. The "slit" used in the present study was actually a small circular aperture with a diameter equal to that of the image of Venus in the image plane of the objective. The resultant spectra had a dispersion of 170 A/mm at 4300 A and a cross-dispersion width of 0·25 mm.

During September and October 1959 we secured 20 exposures on the unilluminated portion of the disk of Venus. The exposure length of these spectra varied from 5 to 80 min on Eastman 103a-F (III), 16 mm film. The spectra obtained show no evidence for the existence of emission lines on the unilluminated portion of the disk.

If we compare these results with those obtained previously, we note the following:

(*1*) The work of Kozyrev resulted in only one satisfactory picture (to quote the author) and this was obtained on the 18th of March 1953. No microphotometer tracings of the region in question (3500–4900 A) were included in the author's paper. The original spectrum may have clearly shown the emission features as reported by Kozyrev, but the existence of these features is not obvious from studies of the published reproduction. The instrument used by Kozyrev consisted of a quartz spectrograph having a dispersion of $\sim$150 A/mm at H_γ used in conjunction with a 50-in. reflector.

(*2*) The work of Newkirk included microphotometer tracings of enlarged positives of the same spectral region as that of the study by Kozyrev. The low dispersion and large film grain made the microphotometer tracings difficult to interpret and left the possibility that an "emission band" as seen on the tracings was actually only a chance superposition of the fluctuations due to grain. Newkirk attempted to express this uncertainty by the probability that a particular emission feature might be the result of such a chance.

(*3*) Short term solar and geomagnetic activity were at about the same level when the spectrograms for all three studies were made although Newkirk's observations in January 1958 were just after the last sunspot maximum. We should, thus, expect the occurrence of "aurorae" on Venus to be equally probable on all three occasions.

(*4*) The reports of visual observation of the "ashen light" by members of the British Astronomical Association are quoted as follows:

> (*1*) "The one really noteworthy feature of the elongation was the unusual prominence of the Ashen Light during the fortnight following December 29 [1957]"[3].
>
> (*2*) ". . . the Ashen Light was not seen at all last autumn [1959]"[4].

In conclusion, we can state that, although the report of spectroscopic emission features during 1958 and their absence in 1959 seems consistent with the visual observations, the existence of emission from the unilluminated side of Venus remains uncertain.

This research was supported under contract 393–05 with the Office of Naval Research.

J. L. WEINBERG
G. NEWKIRK

High Altitude Observatory,
University of Colorado,
Boulder, Colorado

References

1. N. A. Kozyrev, *Izv. Krym. Astr. Obs.*, **12**, 169 (1954).
2. G. A. Newkirk, *Planet. Space Sci.*, **1**, 32 (1959).
3. P. Moore, *J. Brit. Astron. Assoc.*, **69**, 1, 22-26 (1958).
4. P. Moore, *F.R.A.S.*, Private Communication (1960).

Production of Carbon in the Upper Atmosphere*

(*Received* 4 *August* 1960)

1. *Introduction—Earth's Atmosphere*

The lower atmosphere of the earth contains a very small amount[1] (0·033 per cent of volume) of carbon dioxide, mainly supplied by the processes of combustion on the earth. Its content falls with height[2] above the sea-level. At an altitude of about 20 km its volume proportion falls to 0·024 per cent.[3] At higher altitudes, the proportion should fall to vanishing magnitudes.

The main constituents of the earth's lower atmosphere are oxygen and nitrogen which exist in the molecular state in the volume proportions

$$N_2 \sim 70{\cdot}1\%, \quad O_2 \sim 20{\cdot}9\%$$

At higher altitudes, both oxygen and nitrogen dissociate into their atomic form[4]. The density of O_2 molecules decreases with height, at first slowly, and then very rapidly above 100 km. At the same time the density of the atomic oxygen, which is almost zero at 80 km, increases very rapidly with height, attains a maximum at 105 km, and then gradually decreases. Above 110 km oxygen is predominantly atomic.

The situation regarding nitrogen is not so well-established. However, there is a general agreement that nitrogen in the atomic state is significantly present in the upper atmosphere, especially in the auroral regions. An estimate of Bernard[5] that the dissociation of N_2 is one-fifth at the base of the aurora and three-fourths at the upper limit seems very reliable. At very high altitudes, nitrogen should be mostly in the atomic state.

In this technical note, the author intends to suggest a process which may play a significant role in the upper atmosphere. This process involves the interaction of penetrating cosmic rays with the atoms and molecules in the atmosphere. As a result carbon atoms are produced and an equal number of protons released. Because of the high temperatures at these levels, the formation of the oxides of carbon does not seem very likely. However, a few radicals of cyanogen may be formed resulting in a very minute trace of cyanogen gas. Carbon produced in this way is radioactive and decays into nitrogen with the emission of a high energy electron. The protons and electrons released in this process get trapped in the earth's magnetic field. Their contribution to the formation of the inner radiation belt, first observed by Van Allen[6] from observations by the Explorer satellites, may be very significant.

2. *Cosmic Ray Bombardment—Neutron Production*

Penetrating cosmic ray particles, possessing energies as high as 10^{10} eV are able to penetrate deep into the atmosphere. In collisions with the atoms and molecules, they cause disintegrations of the heavier constituents of the upper atmosphere. In the process very fast neutrons are emitted along with other products of disintegration. Since neutrons do not carry an electric charge, they are not subject to the restraint of the earth's magnetic field. Being lighter than the main constituents of the atmosphere, they have a tendency to stay in the upper layers. Hence those emitted in the upward direction follow almost linear paths, and some of them are able to escape out of the atmosphere.

3. *Neutron Decay*

Neutron is an unstable particle and decays freely into a proton and an electron, with the emission of a neutrino

$$n \longrightarrow p + e + \nu \qquad (1)$$

The decay products, protons and electrons, being charged, get trapped in the lines of force of the terrestrial magnetic field, and constitute the inner radiation belt[7-10]. But neutrons have a mean lifetime of about 12·8 min. This

* Originally released as USCEC Rept 56-208, AFOSR TN 59-902 (July, 1959).

is a fairly large period compared to that of mesons which is of the order of 10^{-6} sec, and ensures a good supply of neutrons in the upper atmosphere.

4. *Neutron-Capture*

Before a neutron has a chance to decay, it may collide with a nitrogen atom and cause any one of the following nuclear reactions.

$$ {}_7N^{14} + {}_0n^1 \longrightarrow {}_5B^{11} + {}_2He^4 + Q \qquad (2) $$

and

$$ {}_7N^{14} + {}_0n^1 \longrightarrow {}_6^xC^{14} + p + Q \qquad (3) $$

Only very fast neutrons are effective in these reactions. The first reaction results in the production of a boron atom with the release of an α-particle. The second reaction leads to the formation of atomic carbon with the emission of a proton. Both the boron and carbon atoms have an excess of electrons which they may subsequently drop soon.

5. *Decay of Carbon*

The carbon atoms are radio-active because their nuclei contain 6 protons and 8 neutrons. They decay into the parent nitrogen atoms with the ejection of an electron in accordance with the decay law

$$ {}_6^xC^{14} \longrightarrow {}_7N^{14} + e + Q \qquad (4) $$

In this way atomic nitrogen is recovered. It merely helps the conversion of a neutron into a proton and an electron. The electrons released in this process possess very high energies (1 to 5 MeV).

6. *Carbon—Oxides and Compounds*

The unstable carbon atoms produced in the reaction (3), however, may have a mean lifetime of about 5569 years (very large compared to that of a neutron or a meson). This long life period is sufficient to give a copious supply of carbon atoms in the upper atmosphere. The high temperatures, prevailing at the very high altitudes involved, do not favor the formation or existence of the oxides of carbon (CO or CO_2). However, they may be conducive to the formation of a cyanogen (CN) radical, which, by losing the valence electron of carbon, would readily combine with another (CN) radical to produce a small trace of cyanogen gas $(CN)_2$. It should be of interest to note that no direct evidence of atomic carbon or its compounds has been reported at present. High altitude experiments with probes or satellites should be able to throw some light on this subject.

7. *Radiation Belt*

The protons and electrons—by-products of the mechanism described in the preceding sections—are trapped by the earth's magnetic field. Because of the abundant supply of atomic nitrogen and large number of available fast neutrons at high altitudes, this process should be of significance in the formation of the inner radiation belt. The importance of the neutron-decay theory has been very much emphasized recently[7–10]. This process is akin to the neutron-decay process in that a neutron converts itself into a proton and an electron—only it has to undergo a little more complicated routine. In the author's opinion, both the neutron-decay and neutron-capture processes are in strong competition with each other in the formation of the inner radiation belt.

8. *Conclusions and Discussions*

(*a*) Carbon and boron atoms should be present in the upper atmosphere. A small trace of cyanogen gas should also be expected.

(*b*) Neutron-capture by nitrogen atoms should make a significant contribution to the formation of the inner radiation belt.

In order to evaluate and assess the predictions of this hypothesis, one needs to have an exact knowledge of the number of neutrons and nitrogen atoms per c.c. available at high altitudes, the energy spectrum of neutrons, the prevalent temperatures, the cross-section of collisions between neutrons and atomic nitrogen, and the relative efficiencies of the neutron-capture reactions. We plan to come back to these questions in a more detailed subsequent note. We will, however, show in a relatively simple way here that the probabilities of neutron-capture and neutron-decay are of the same order.

In order to make the comparison easy, we adopt the same scale length of 100 miles (160 km) as used by Singer[7] in his estimates. We assume that the neutrons have an average speed of 10^{10} cm/sec. Therefore the time interval t required, on the average, to describe a distance of 160 km equals $1{\cdot}6 \times 10^{-3}$ sec. Since the mean lifetime of free neutrons is 780 sec, the probability of a free neutron decay within a path-length of 160 km is approximately given by

$$p_d \sim \frac{t}{T} \sim 2 \text{ in } 1 \text{ million}$$

On the other hand, if we assume that each collision results in neutron capture and that the particle density of atomic nitrogen at an altitude of 200 km is 10^{10}/c.c., the probability of neutron capture within a path-length of 160-km is given by

$$p_c \sim \sigma n s \sim 1 \text{ in } 2 \text{ million}$$

This is a very conservative estimate but certainly leads to the significance of the neutron capture in the upper strata of our atmosphere.

K. P. CHOPRA

Polytechnic Institute of Brooklyn,
Aerodynamics Laboratory,
527 Atlantic Avenue,
Freeport, New York

REFERENCES

1. G. P. KUIPER, *The Earth as a Planet*. Chapter 7, Chicago University Press, Chicago (1954).
2. G. P. KUIPER, ibid, Chapter 8.
3. M. SHEPHERD, *Nat. Geog. Soc. Contr. Tech. Pap., Stratosphere Ser.* **2**, 117 (1936).
4. S. K. MITRA, *The Upper Atmosphere*. Chapter V, Asiatic Society, Calcutta (1952).
5. R. BERNARD, *Phys. Rev.* **55**, 511 (1939). *Ann. d'ap.*, **4**, 13 (1941).
6. J. A. VAN ALLEN, G. H. LUDWIG, E. C. RAY and C. E. MCILWAIN, I.G.Y. Satellite Rept. No. 3, (1958). *Jet Propulsion*, **28**, 588 (1958).
7. S. F. SINGER, *Bull. Am. Phys. Soc.*, **1**, 229 (1956). *Phys. Rev. Letters*, **1**, 171 (1958); *J. Brit. Interplan. Soc.*, **16**, 558 (1958).
8. S. N. VERNON, *CSAGI Meeting*, Moscow (1958).
9. P. B. KELLOG, *Nuovo Cimento*, (1959).
10. N. C. CHRISTOFILOS, *UCRL Rept.* 5047 (1958).

The Microwave Temperature of Venus*

(*Received* 16 *December* 1960)

The measurements of the brightness temperature of Venus at or near several inferior conjunctions have indicated a rather high brightness temperature of about 600°K in the centimeter region. The values at 3 and 10 cm[1–5] apparently indicate that the centimeter radiation follows a black-body curve and it is generally assumed that this comes from a region at or near the surface of the planet. The measurements in the millimeter region by Gibson and McEwan[6] and by Kuzmin and Salomonovich[7] show a decrease in brightness temperature of 200–300°K from that at 3 and 10 cm which might be interpreted as indicating a cooler cytherean atmosphere which absorbs the microwave energy coming from the surface. Because it is known that a great amount of carbon dioxide exists above the cloud layer and because of the recent detection of water vapor[8] there can be considered to be an atmosphere that might be capable of fulfilling this requirement.

It is immediately obvious that to account for such a hot surface a very efficient greenhouse effect must be assumed. Sagan[9] has analyzed the available water vapor and CO_2 data and concluded that such a greenhouse effect could indeed exist.

Barrett[10] has attempted to fit the microwave temperature data by calculating the brightness temperature dependence on wavelength in the critical millimeter and short centimeter regions. Various atmospheric models were assumed for the region below the clouds. In order to fit the 8 mm and centimeter data it is apparent that the surface pressure must be greater than 10 atm with 3 per cent H_2O or 10 atm with greater H_2O. The amount of water vapor necessary to sustain the required greenhouse effect indicates that, in order to fit the data, the surface pressure must be much higher than the value of several atmospheres one obtains from the commonly accepted values of pressure and temperature at the cloud tops of roughly

* This paper presents results of research carried out under Contract No. NASw-6, sponsored by the National Aeronautics and Space Administration.

1/6 atm and 285°K respectively. Barrett has not included the attenuation due to the clouds themselves, but it appears that the amount of liquid water content one must assume is too great considering Strong's data, and the solid particles considered to be in the cloud layer are apparently of insufficient size, etc.

One possible mechanism which could cause the high brightness temperature in the centimeter region and predict a lower millimeter temperature is that of free-free emission by electrons in an optically thick cytherean ionosphere. We have, for an isothermal layer at an electron temperature T_e, an effective temperature

$$T_b = T_e(1 - e^{-\Delta}) + T_s e^{-\Delta}$$

where Δ is the opacity as given by (for $Z=1$, $n_e = n_i$)

$$\Delta = \frac{1{\cdot}98 \times 10^{-23} g\lambda^2}{T_e^{3/2}} \int n_e^2 dz$$

where g is the Gaunt factor (taken to be ~ 4 here). If one assumes an electron temperature of ~ 600°K and an average surface temperature of $T_s \sim 265$°K, although the latter can be as high as perhaps 380°K, a fairly good fit to the data is obtained if one assumes

$$\int n_e^2 dz \sim 10^{25}/\text{cm}^5.$$

The possibility that part of the 3 cm radiation might come from a heavily ionized layer of the atmosphere has previously been considered[1].

If Venus has a magnetic field that is comparable to or greater than Earth's, it is difficult to see how such an ionized layer can be maintained against recombinations. However, if the field is small, say 1/30th that of the Earth, then one finds a tempting mechanism in the solar corpuscular radiation which consists of electrons and protons streaming out from the Sun. Taking $\int n_e^2 dz \sim 10^{25}/\text{cm}^5$, we have (for oxygen at 600°K)

$$\alpha \int n_e^2 dz \sim 10^{13}/\text{cm}^2 \text{ sec}$$

If we have $\sim 10^3$ protons/cm^3 streaming into the outer reaches of the cytherean atmosphere at ~ 750 km/sec then the proton flux is roughly $7{\cdot}5 \times 10^{10}$ protons/cm^2 sec. Since such protons have energies in the order of kilovolts we see that such an ionosphere could be sustained. It should be mentioned that these particle densities and velocities are not inconsistent with what one might expect at the orbit of Venus assuming some of the prevailing solar wind theories.

It is interesting to note that if one allows any or all of the variables T_s, T_e and n_e to be functions of the zenith angle, a fit to about any variation in T_b with phase observed can be made. As an example, if one assumes that the electron and surface temperature do not vary with zenith angle, etc., but that the electron density does, in a manner similar to that on Earth, then a crude fit to the 1959 8 mm and 3·75 cm phase data can be obtained for the case where the noontime electron density is about 3 to 5 times that at night. In their interpretation of the variation in the 8 mm brightness temperature with phase angle, Kuzmin and Salomonovich[7] have apparently neglected this possibility.

D. E. JONES

Jet Propulsion Laboratory,
California Institute of Technology,
Pasadena,
California

References

1. C. H. MAYER, T. P. MCCULLOUGH, and R. M. SLOANAKER, *Astrophys. J.* **127,** 1 (1957).
2. C. H. MAYER, *Astr. J.*, **64,** 43 (1959).
3. C. H. MAYER, T. P. MCCULLOUGH, and R. M. SLOANAKER. Paper presented to the XIIIth General Assembly, URSI, London, 5-15 Sept. 1960.
4. L. E. ALSOP, J. A. GIORDMAINE, C. H. MAYER, and C. H. TOWNES, *Astr. J.*, **63,** 301 (1958).
5. F. DRAKE. Private communication.
6. J. E. GIBSON and R. J. MCEWAN, *Proc. IAU Symp. No. 9-URSI Symp. No. 1* (ed. R. N. Bracewell), University Press, Stanford.
7. A. D. KUZMIN and A. E. SALOMONOVITCH, *Astron. Zh.* **37,** 297 (1960).
8. J. STRONG. Private communication to R. W. Davies, Jet Propulsion Laboratory (1960).
9. C. SAGAN. Unpublished Ph.D. Dissertation, Yerkes Observatory, University of Chicago (1960).
10. A. H. BARRETT, *J. Geophys. Res.*, **65,** 1835 (1960).

PAPERS TO BE PUBLISHED IN THE NEXT ISSUE

TRANSLATED IN FULL FROM

ARTIFICIAL EARTH SATELLITES U.S.S.R.
(ISKUSSTVENNYE SPUTNIKI ZEMLI)

I. S. SHKLOVSKII: Optical methods of observation of artificial earth satellites.

L. I. SEDOV: Dynamic effects in the motion of artificial earth satellites.

V. I. KRASOVSKII: Soviet investigation of the ionosphere with the aid of rockets and artificial earth satellites.

S. SH. DOLGINOV, L. N. ZHUZGOV and N. V. PUSHKOV: A preliminary report on geomagnetic measurements carried out on the third Soviet artificial earth satellite.

V. I. KRASOVSKII, IU. M. KUSHNIR, G. A. BORDOVSKII, G. F. ZAKHAROV and E. M. SVELITSKII: The detection of corpuscles with the aid of the third Soviet artificial earth satellite.

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 169–184. Printed in Great Britain.

ATMOSPHERIC PHENOMENA NOTED IN SIMULTANEOUS OBSERVATIONS OF 1958 δ2 (SPUTNIK III)

J. AARONS and H. E. WHITNEY
Air Force Research Division, United States

R. S. ROGER and J. THOMSON
Jodrell Bank Experimental Station, England

J. BOURNAZEL and E. VASSY
Laboratory of Atmospheric Physics, France

H. A. HESS and K. RAWER
Ionospheric Institute of Breisach, Western Germany

N. CARRARA and P. CHECCACCI
Centro Microonde, Italy

B. LANDMARK and J. TRÖIM
Norwegian Defense Research Establishment, Norway

B. HULTQVIST and L. LISZKA
Kiruna Geophysical Observatory, Sweden

(*Received* 23 *August* 1960)

Abstract—Amplitude and doppler records of the 20 Mc/s transmissions of 1958 δ2 (Sputnik III) were simultaneously made at six radio observatories in Europe. Several distinct phenomena were noted.

A sharp decrease in amplitude, which lasted from ½–1½ min and was followed by an abrupt return of the signal, occurred at the same time at all stations. This phenomenon has been termed "dropout", although indications of low-level signals were usually present, and is explained as a sudden decrease in transmitter power of the satellite perhaps when the satellite passed through a region of high electron flux.

Another type of decrease in the signal was attributed to highly localized absorbing or scattering regions.

A third phenomenon, recorded by techniques of simultaneous measurements, was observed in the concurrence of scintillating signals at some stations and Faraday fading at others. The overlap of the propagation path to each station recording the scintillations defines a vertical slab of ionospheric region responsible for the scintillations. The region is about 400–600 km in the horizontal plane but includes small scale irregularities of less than 1 km.

In a joint study participated in by a group of six radio observatories in Europe and one in the United States, amplitude and doppler recordings of the transmissions from 1958δ2 (Sputnik III) were made in the 3-week period that included the date of the solar eclipse on 2 October 1959. From the data examined at a subsequent conference of the group, three revolutions illustrating atmospheric effects typical of the series have been chosen for discussion in this report.

The laboratories making their observations, the co-ordinates of the field sites and the antenna configuration used are shown in Table 1.

The stations differed somewhat in their techniques of observation and means of recording the data. For the amplitude measurements, all used relatively wideband stable-frequency receivers, to feed d.c. amplifiers and recording milliameters but the doppler techniques varied. Val-Joyeux, for example, obtained its doppler data by observing lissajous figures displayed on an oscilloscope from the magnetic tape recording. Although slight inaccuracies in synchronization may limit the total precision of

Table 1. Geographical Co-ordinates of the Satellite Stations and Antenna Orientation

	Long. (E)	Lat. (N)
Jodrell Bank Experimental Station University of Manchester, England λ/2 folded dipole N–S	357° 41′ 47″	53° 14′ 01·7″
Norwegian Defense Research Establishment Kjeller, Norway λ/2 folded dipole N–S	10° 46′	60° 13′
Kiruna Geophysical Observatory Kiruna, Sweden λ/2 dipole E–W	20° 26′	67° 50′
Ionospheric Institute of Breisach, Breisach, W. Germany λ/2 dipole E–W	7° 35′ 24″	48° 03′ 18″
Laboratory of Atmospheric Physics Paris, France (Station at Val-Joyeux) λ/2 dipole N–S	2° 1′	48° 49′
Centro Microonde Florence, Italy λ/2 folded dipole N–S	11° 13′ 53″	43° 48′ 11″

the sets of measurements, they are not enough to affect the overall results.

In making initial comparisons of the amplitude tracings, it was noted that Faraday fades were being received at some stations while scintillating signals were being received at others. The change in character and disturbed appearance of the signals has been ascribed to properties of the propagation path to the stations receiving these scintillations. Comparative time graphs of the starts and stops of the scintillating effects recorded at the various stations maps the dimensions of scintillation-producing regions in the ionosphere. By extension, the height of the disturbed region can be estimated.

1. REVOLUTION 7172 (3 OCTOBER 1959)

1.1. *Path of the satellite*

The recordings of revolution 7172 taken at the European stations on the morning of 3 October 1959 best illustrate the data gathered and analysed. The path of the satellite is shown in Fig. 1. The geomagnetic activity index K_p for the 3-hr interval 0300–0600 was 3. The North Atlantic CRPL Radio Propagation quality figure for the 6-hr period was 6, indicating an average transmission period. Ionograms at the various stations were similar.

1.2. *Localized ionospheric regions*

Signals were detected at $4^h24^m22^s$ by the Laboratory of Atmospheric Physics in Val-Joyeux, France. This group made only doppler observations for the period. At the same time, however, the Centro Microonde in Florence noted a signal level increase on its amplitude recordings. Fig. 2 shows 2 minutes of the Florence tracing, from 4^h28^m when maximum amplitude of the signal was recorded, through the period when the signal level fell below the noise, to 4^h30^m when the amplitude once again increased. Clear Faraday fading was noted on the Florence recording.

On the other hand, Jodrell Bank records from $4^h26^m30^s$ to 4^h29^m (Fig. 3) were characterized by scintillations. This is clear indication that the ionosphere between the satellite and Jodrell Bank was quite different from the ionosphere between the satellite and Florence.

It should be emphasized that at times in this study the distinction between scintillations and Faraday fading is necessarily somewhat objective. The criterion followed is that Faraday fades display a neat maximum and minimum for several pulses of the transmitter, probably best illustrated in the Florence record (Fig. 2). In revolution 7172 the periods of Faraday fading at 20 Mc/s vary from 5 to 20 sec. Since the individual satellite pulses are at the rate of 3 pulses/sec each Faraday rotation encompasses more than one satellite pulse. Scintillations rarely show clean minima for several pulses, nor do the periods between their maxima exhibit the relatively slow change over several minutes seen in Faraday fading. A mixture of the two types of signals, called *Faraday-scintillations*, occurs when the propagation path is close to the boundary of a scintillation region.

In the recordings of Jodrell Bank (Fig. 3), scintillations are evident in the time period before 4^h29^m. After $4^h29^m30^s$, Jodrell Bank shows Faraday fading following low-amplitude Faraday fading that begins at $4^h28^m50^s$.

Fig. 1. Path of revolution 7172 across Europe on Oct. 3, 1959.

At the Kjeller station of the Norwegian Defense Research Establishment, fast-run recordings for this period showed clear Faraday fading in low-amplitude signals received shortly after 4^h27^m.

The data thus indicates that ionospheric regions producing scintillations are relatively localized—Jodrell Bank recorded scintillating signals, but Florence and Kjeller did not. Such a region may have caused the unusual change in doppler recorded at Val-Joyeux from $4^h25^m30^s$ to $4^h26^m30^s$. The doppler curve of the pass at

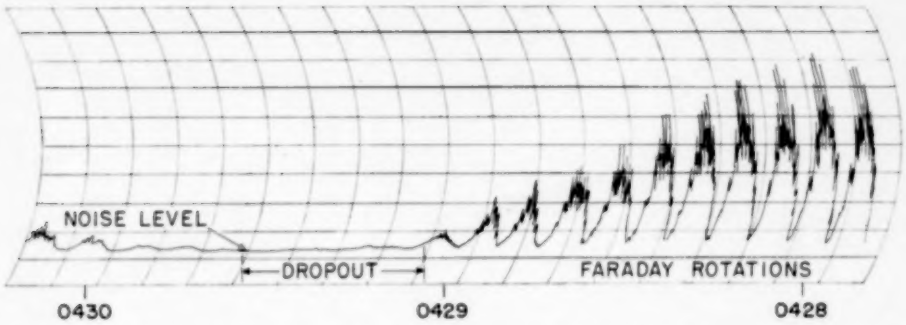


Fig. 2. The clean Faraday rotations noted are characteristic of the signals observed at this latitude on almost all high altitude passes.

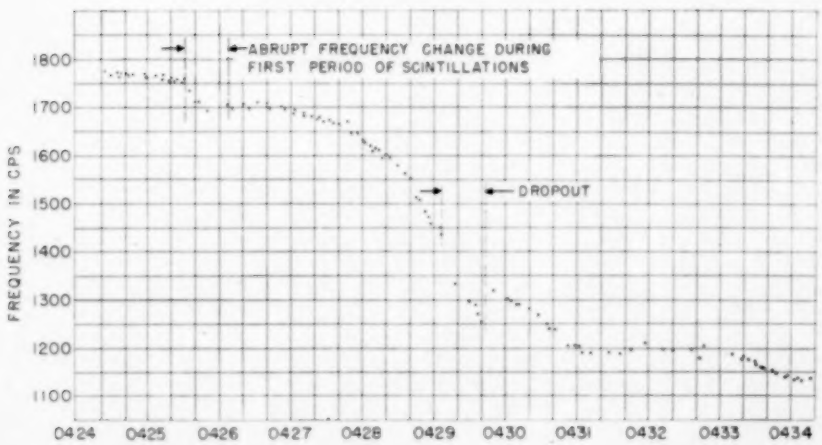


Fig. 4. Val-Joyeux, France, doppler data on 1958 δ2, revolution 7172.

(Facing page 172)

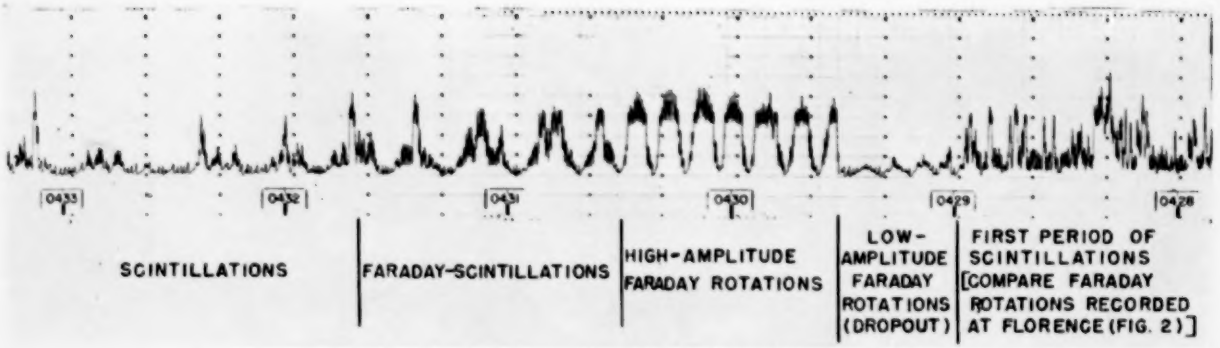


Fig. 3. Jodrell Bank recordings of 1958 $\delta 2$ revolution 7172.

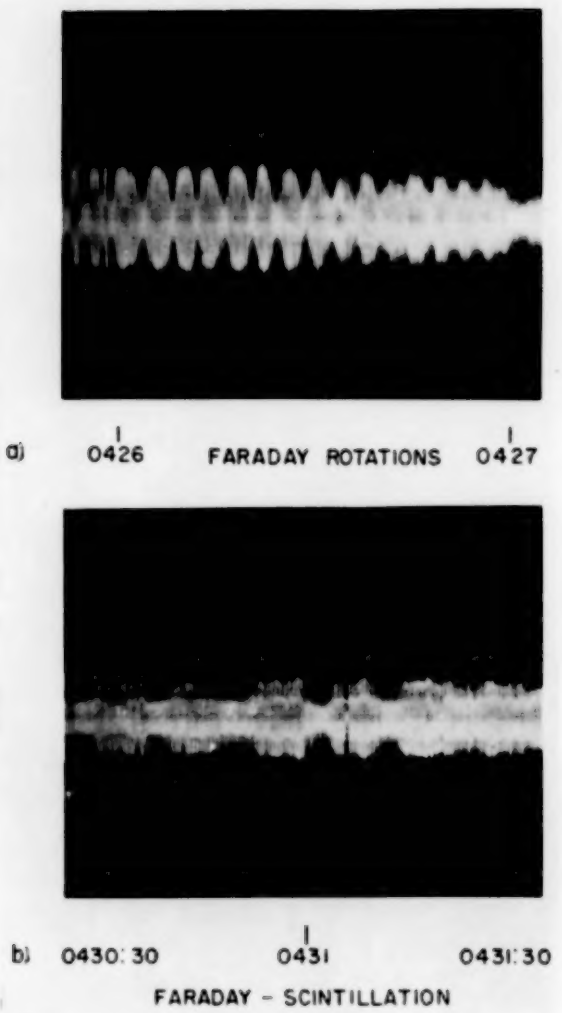


Fig. 8. 20 Mc/s amplitude recordings of Breisach, Germany, of revolution 7172.

Val-Joyeux (Fig. 4) shows a sharp decrease in frequency at the beginning of this one-min period.

The region producing the scintillations could not have been too close to the satellite or Kjeller signals would also have shown scintillation.

1.3. "*Dropout*"

The second event in the pass was signal "dropout," a sudden decrease in signal recorded at all stations simultaneously. The signal did not drop to noise level at all stations; low-level signals were still observed at some. Reception increased just as abruptly as it decreased, also simultaneously at all stations. The Florence and Jodrell Bank records (Figs. 2 and 3) show this dropout period. Breisach 40 Mc/s observations confirmed the dropout. Inadvertent errors in timing caused the slight difference beween the records.

The Kjeller fast-run records (Figs. 5(*a*)–(*d*)) show pulse-to-pulse changes in shape produced by atmospheric conditions as well as by the satellite transmitter. The normal pulse shape, its flat-top form and the Faraday fading are shown in Fig. 5(a). Shortly after $4^h28^m33^s$ the form of the received pulse changes to a sloping top (Fig. 5(b)). Then (Fig. 5(c)), for 30 sec from $4^h29^m7^s$, the start of the dropout period, there occurs an unusual modulation of the transmitter pulse. That the transmitter characteristics are changing is evidenced in the double pulses that now replace every alternate pulse, a sort of splitting of the wider pulses seen in Fig. 5(a) and (b). It is presumed that for this short period the change in modulation decreased transmitter power from the satellite and thus caused a low signal level or dropout. (Fig. 5(d) is discussed in Section 1.5, which deals with scintillation.)

Drops in signal amplitude to near noise level

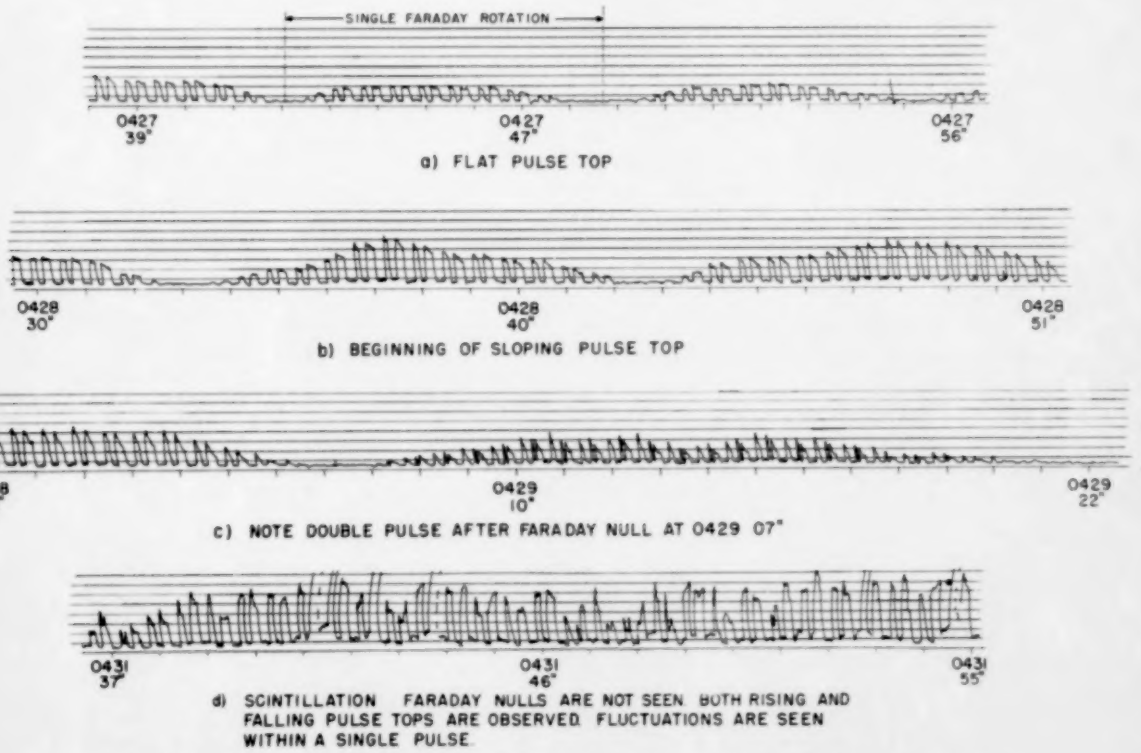

SINGLE FARADAY ROTATION
0427 39"
0427 47"
0427 56"
a) FLAT PULSE TOP
0428 30"
0428 40"
0428 51"
b) BEGINNING OF SLOPING PULSE TOP
0428 59"
0429 10"
0429 22"
c) NOTE DOUBLE PULSE AFTER FARADAY NULL AT 0429 07"
0431 37"
0431 46"
0431 55"
d) SCINTILLATION FARADAY NULLS ARE NOT SEEN BOTH RISING AND FALLING PULSE TOPS ARE OBSERVED FLUCTUATIONS ARE SEEN WITHIN A SINGLE PULSE

Fig. 5. Kjeller, Norway, fast run recordings of 1958 δ2, revolution 7172. Logarithmic amplitude scale.

that were recorded simultaneously at all stations were accompanied by changes in the received doppler frequency. The simultaneity of the dropouts and the abruptness of the frequency changes suggest that the satellite apparatus is involved. In the doppler data of Val-Joyeux (Fig. 4) an abrupt drop of about 75 c/s occurs between 4^h29^m and $4^h29^m32^s$. Discontinuous jumps of up to 100 c/s, usually lasting less than one minute, were common in the measured frequency of Sputnik III. The Cavendish Laboratory, Cambridge, Nicholl[1], observed that changes similar to these at the 20 Mc/s frequency received from Sputnik III were matched by proportional frequency changes in the second harmonic at 40 Mc/s. Any shift due to ionospheric causes would of course be less severe at 40 Mc/s than at 20 Mc/s.

In the Val-Joyeux dropout the frequency jump was probably the fault of the satellite transmitter. Abrupt frequency changes without concomitant variations in modulation are more likely caused by ionospheric conditions since it is unlikely that the transmitter should suddenly change frequency. If the frequency jumps were of fixed value and periodic they might be caused by the roll of the satellite body or the effect of sunlight on one area of the satellite. Non-periodic changes in frequency or amplitude are most likely related to conditions external to the transmitter.

It may be that dropout is a temporary breakdown induced by high-energy particles in the satellite transmitter or antenna. What is more likely, the hypothesis elaborated in Section 4, is that the change in transmitter modulation and lowering of output power is induced by the attempt to count high-energy particles in the satellite.

1.4. *Amplitude measurements*

Figure 6 presents signal amplitudes recorded at five stations, collated on a relative scale. The Kjeller, Kiruna, Florence and Jodrell Bank data is for 20 Mc/s. The Breisach data is for 40 Mc/s. Each Faraday fade, evaluated by the station recording it, is marked with a dot at its peak amplitude. The mean value of the scintillations is marked by crosses. At Kiruna, where clear Faraday fading is rarely seen as it is at Florence, the peak values of both scintillation and Faraday fading are used. It is difficult, however, without precise knowledge of both the ground and satellite antenna orientations, to draw any general conclusions on absorption from this simple summing-up of the amplitudes.

1.5. *Second scintillation period*

In the 20 Mc/s recordings after modulation of the pulse had stopped and the signal level had returned to its higher value, various starts and stops of scintillations were observed. This data is listed in Table 2 and also translated into graphic form on the map in Fig. 7.

Table 2

Florence, Italy	Faraday rotations $4^h29^m30^s$ to 4^h32^m
Jodrell Bank, England	Faraday rotations $4^h29^m30^s$ to $4^h30^m30^s$
	Faraday and scintillations $4^h30^m30^s$ to $4^h31^m45^s$
	Scintillations $4^h31^m45^s$ to 4^h33^m
Kjeller, Norway	Scintillations from $4^h30^m45^s$ to 4^h32^m
Kiruna, Sweden	Scintillations from $4^h29^m30^s$ to $4^h31^m15^s$

The recordings of Faraday rotations in Fig. 8(a) and the scintillations in Fig. 8(b) were made by Breisach at 20 Mc/s. Breisach also recorded at 40 Mc/s, but the chosen technique made it difficult to perceive the start and stop of Faraday fading and scintillations. That the irregular fluctuations extend to 40 Mc/s can, however, be discerned.

If the propagation path during the period in which the scintillations are observed at each station is joined to the subsatellite points, the scintillating region in the ionosphere is defined as shown in Fig. 7. The height of the satellite at the time shown was approximately 1000 km. By taking the point in the intersected region

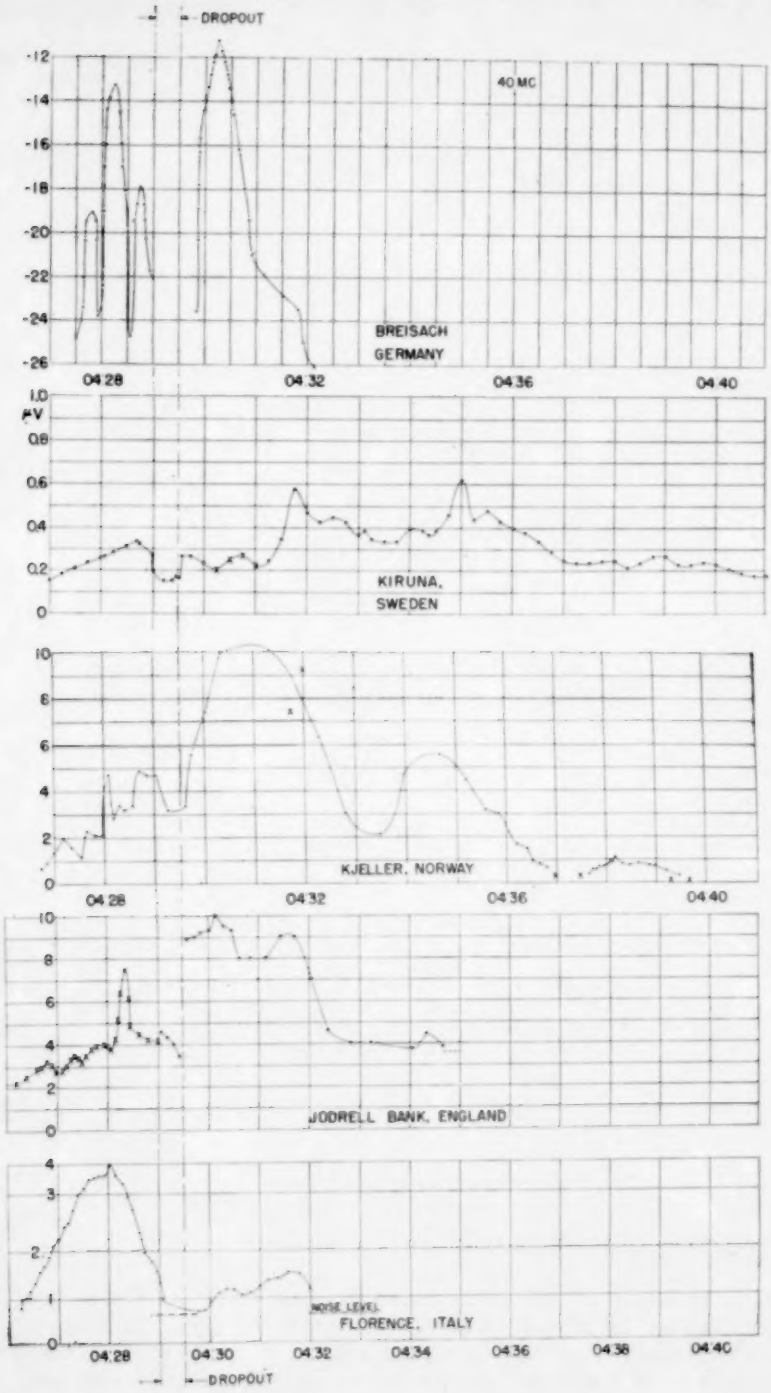


Fig. 6. Collation of amplitude on uniform time scale. Each dot represents the peak value of a single Faraday rotation while a cross is used on the mean value of the scintillation amplitude. An arbitrary amplitude scale is used.

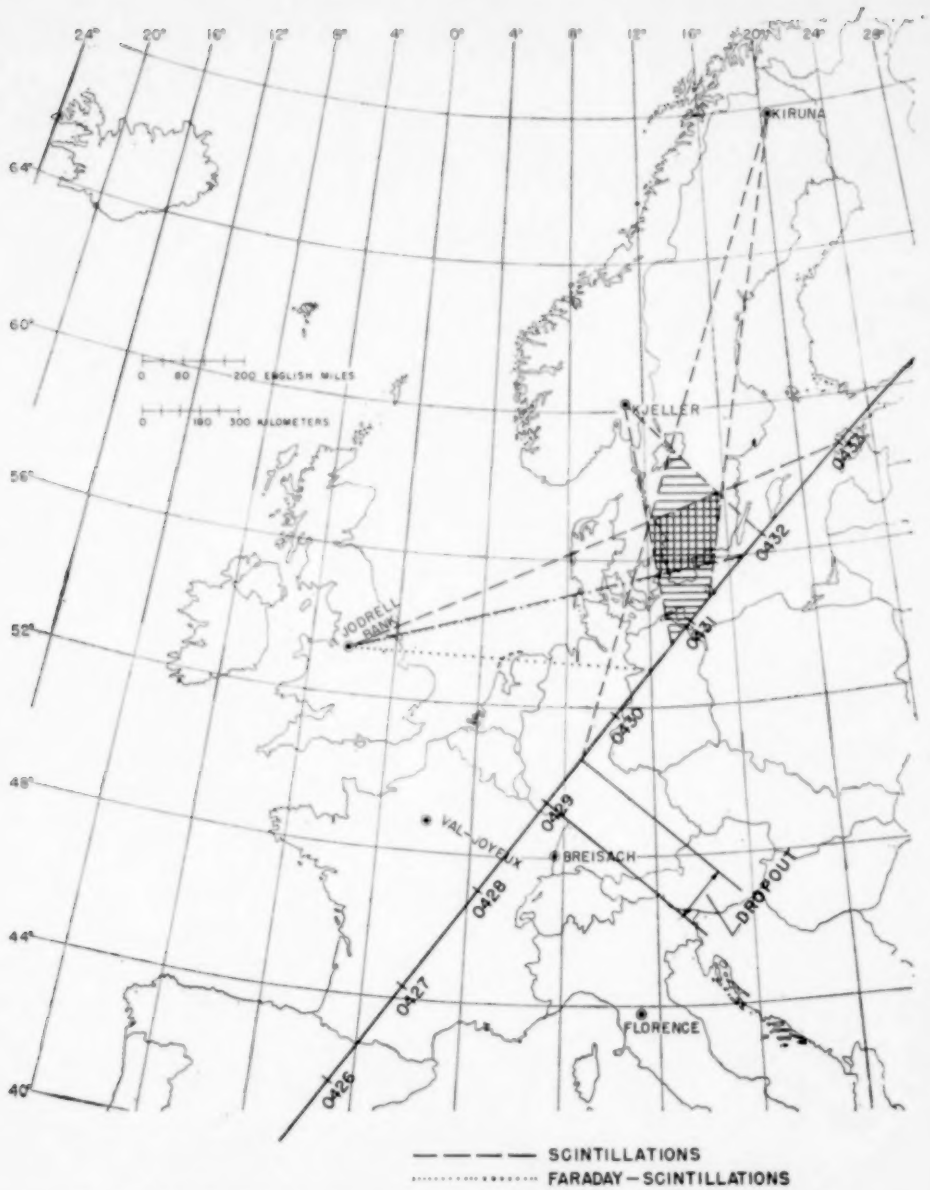


Fig. 7. Scintillation region observed during 1958 δ2 revolution 7172, 3 October, 1959

closest to Kjeller at 4^h32^m, the lowest altitude of the scintillating region can be calculated. The approximate value is 250 km. Greater precision is unwarranted because it is difficult to estimate the start and stop of scintillations. The 4^h31^m observations of scintillations at Kiruna, Kjeller and Breisach (and Faraday-scintillations at Jodrell Bank) indicate that the

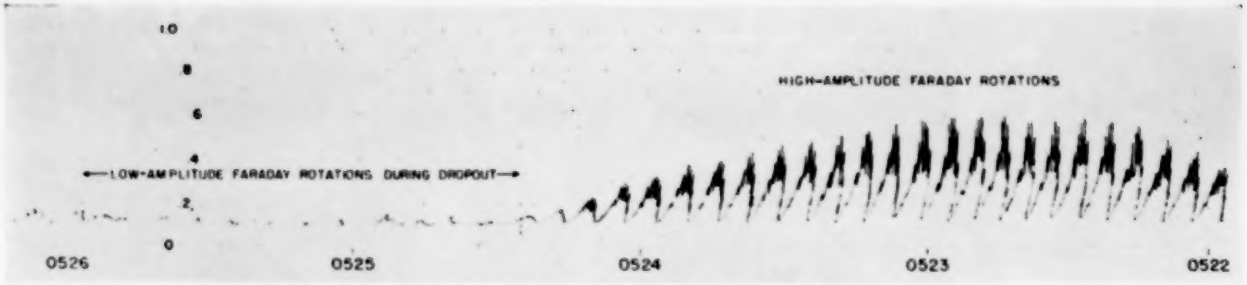


Fig. 11. Florence, Italy, recordings of 1958 δ2, revolution 7128, Sept. 30, 1959.

(To face page 176. Refers to page 178)

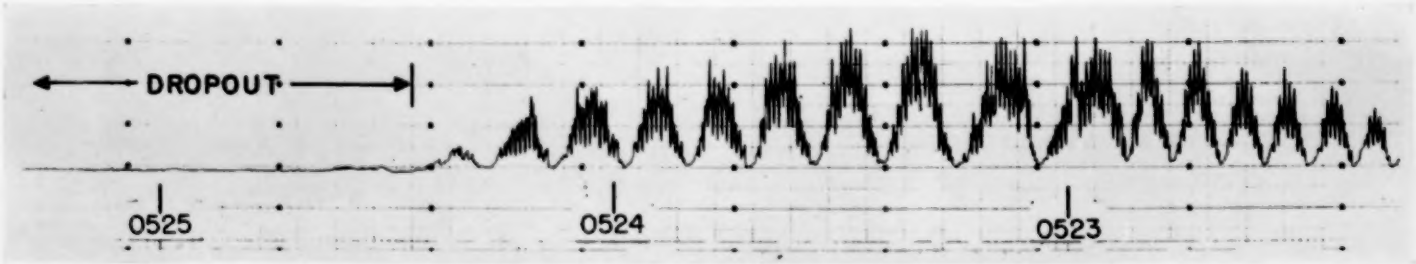


Fig. 12. Jodrell Bank recordings of 1958 δ2, revolution 7128, Sept. 30, 1959.

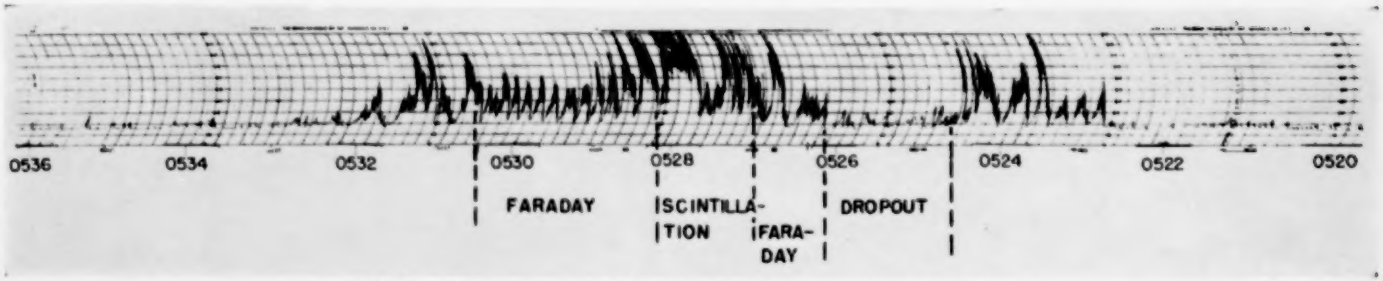


Fig. 13. Kjeller recordings of Faraday rotation, dropout and scintillation.

region may indeed extend to the height of the satellite.

From the configuration of the intersection of the paths of the stations, it is apparent that a roughly elliptic (major axis N–S) localized region in the ionosphere is responsible for scintillations. Its height may range from 250 to 1000 km, a value that receives confirmation from the Jodrell Bank data. While scintillations were being received at Kiruna and Kjeller, Jodrell Bank was receiving a mixture of Faraday fading and scintillation, the earlier clear Faraday fades having changed to the combination signals with secondary minima (Fig. 3). At this time the propagation path to Jodrell Bank was along the minor axis of the slab responsible for the scintillations, whereas the propagation path to the other stations was along the major axis of the ellipse.

It should be noted that in the first period of scintillations (4^h27^m–4^h29^m) when Jodrell Bank was receiving scintillations—and Breisach, as well as Florence and Kjeller, were receiving Faraday fading—the propagation path to Jodrell Bank was to the north, through the major-axis long-pathlength portion of the ionospheric slab.

Figure 5(d), the Kjeller data (see Section 1.3) indicate the effect of the scintillation region on the pulse shape. The normal, relatively flat, pulse form seen in Fig. 5(a) is so changed by the region that within the pulse there can be observed fast amplitude fluctuations indicating the rapidly changing propagation path through which the propagated pulse is passing. Since the path irregularities vary considerably even across the distance the satellite travels in 0·1 sec, they are probably less than 1 km, horizontally massed together in a larger region extending for 600 km.

Observations in the northern latitudes have always disclosed the presence of irregularities. Kiruna registers scintillations on many passes. Even during revolution 7172 (Fig. 9), several periods of scintillations and mixtures of Faraday-scintillations were recorded. The constant rate of change of Faraday rotation almost always evident at Florence is rarely seen at Kiruna. Thus, the Kiruna records show the many discontinuities in electron density and possibly in the magnetic field.

2. REVOLUTION 7128 (30 SEPTEMBER 1959)

The phenomena analysed in revolution 7172 are also illustrated in revolution 7128, which the stations recorded between 5^h19^m and 5^h35^m on 30 September 1959. The time period had a *K* index of 1+. The height of the satellite ranged from 1031·5 km at 5^h19^m to 910 km at 5^h35^m. The subsatellite points are plotted in Fig. 10.

In analysing the reduced data and original amplitude recordings, two distinct phenomena

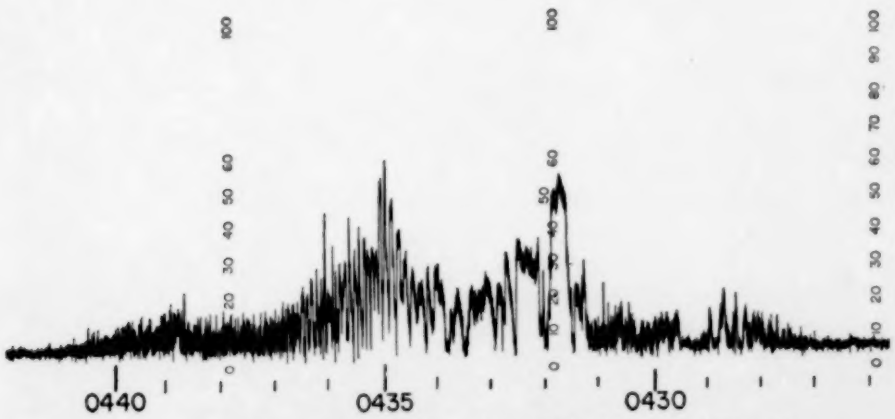


Fig. 9. Kiruna observations of 1958 δ2 revolution 7172, 3 October 1959. (Note irregular periods of Faraday rotation in contrast to Florence observations.)

Fig. 10. Path of revolution 7128 across Europe on Sept. 30, 1959.

are again noted: (1) all stations simultaneously recorded a dropout that lasted 1½ min; (2) scintillating signals were recorded at one or more stations while Faraday fading was recorded at others.

The signals recorded at Florence (Fig. 11) appeared at about 5^h20^m and increased in amplitude until $5^h22^m30^s$. They then decreased, cutting out almost completely at $5^h24^m30^s$, after which low-level Faraday fading occurred until higher-amplitude signals returned at 5^h26^m.

The Jodrell Bank amplitude recordings (Fig. 12) are quite similar to the Florence data, but the Kjeller records (Fig. 13) are not. By

5^h24^m the character of the Norwegian signals had changed from Faraday fading to an erratic fluctuation. A sharp cutoff was observed at the time of the dropout, at $5^h25^m30^s$.

The Breisach records for $5^h23^m45^s$ show a behaviour similar to Faraday-scintillations. Cutoff was observed at $5^h25^m30^s$ at Breisach, as it was at Val-Joyeux, with a return to higher signal amplitudes $1\frac{1}{2}$ minutes later. In both revolutions, 7128 and 7172, dropout occurred at about 50° latitude. During 7128, when the altitude of the satellite was approximately 1000 km, the subsatellite path was 425 km in extent.

With the return of a high signal level, there was a period when some stations detected scintillating signals at the same time that others detected slow Faraday fading. (See Table 3.)

Table 3. Signals at Four Stations from $5^h26^m20^s$–5^h28^m

	$5^h26^m20^s$ to 5^h27^m	5^h27^m to $5^h27^m40^s$	$5^h27^m40^s$ to 5^h28^m
Jodrell Bank	Faraday	Scintillation	Faraday-scintillation
Kiruna	Scintillation	Faraday from $5^h27^m10^s$ on	—
Kjeller	Faraday	Scintillation	Scintillation
Breisach	Faraday-scintillation on 40 Mc/s (times not inclusive)	—	—

Figure 14 shows the periods of scintillation for each station marked on an expanded version of the map of the satellite's track. Overlap of the lines drawn from the satellite to the stations indicates that a disturbed ionospheric region is being crossed by signals transmitted from the satellite. The geometry shows that the maximum height for this region is rather high, perhaps 1000 km; the data indicates that the lower height is 400 km.

Although the general geographic extent of this region is rather clearly defined, there is a basis for some difference of opinion mainly because the Faraday rotations cannot always be positively distinguished from the scintillations. A slight north–south elongation is apparently indicated, but the data from the two revolutions is not conclusive on this point.

Kiruna amplitude recordings (Fig. 15) are characterized by clear Faraday fading only between 5^h29^m and $5^h32^m30^s$; before 5^h29^m there is a mixture of scintillation and Faraday fading.

Val-Joyeux records (Fig. 16) indicates a slight change of frequency (about 25 c/s) at $5^h22^m30^s$ and during the dropout period.

The Breisach 7128 data at 40 Mc/s, together with the data at 20 Mc/s for all other stations except Val-Joyeux, is collated in Fig. 17.

3. LOCALIZED REGION OF HIGH IONIZATION

An effect that could be confused with dropout if it appeared at only one station was noted on 25 Sept. 1959 during revolution 7055 of 1958 δ2. The Kjeller, Jodrell Bank, Florence and Kiruna records for this pass are reproduced in Fig. 18. At 7^h22^m the subsatellite point was almost directly east of Jodrell Bank, on the coast of England. In the south-north passage at this moment, Kjeller recorded a sudden sharp decrease in signal that was just as abruptly followed by a sudden return to a high signal level approximately 15 sec later. At this time neither Jodrell Bank nor Florence showed any such decrease in signal intensity. Kiruna showed a slight decrease from its rather low level. From the appearance of the records at Kjeller alone, the phenomenon might appear to be dropout, yet a comparison of the data at all stations indicates the existence of a region more highly ionized than its surroundings. The data is too meager to determine whether the sudden decrease in the signal is due to absorption or scattering.

4. DISCUSSION

4.1. *Sharp decrease of signal level*

The sudden decreases observed in the satellite signal are apparently due to two different causes: (1) local absorption or scattering (see Kjeller, Fig. 18) is indubitably ionospheric in origin; (2) the dropout phenomenon must be a function of the satellite transmitter and perhaps of the satellite antenna.

Dropout recordings show the persistence of

Fig. 14. Scintillation region observed during revolution 7128.

low-level signals with, in some instances, Faraday fading. Because of the pulse modulation, the continuation of Faraday rotations, the sharpness of the effect, and because it would have to occur at the altitude of the satellite itself, dropout cannot be explained as a propagation effect. If a mechanism similar to other known processes of ionospheric absorption is

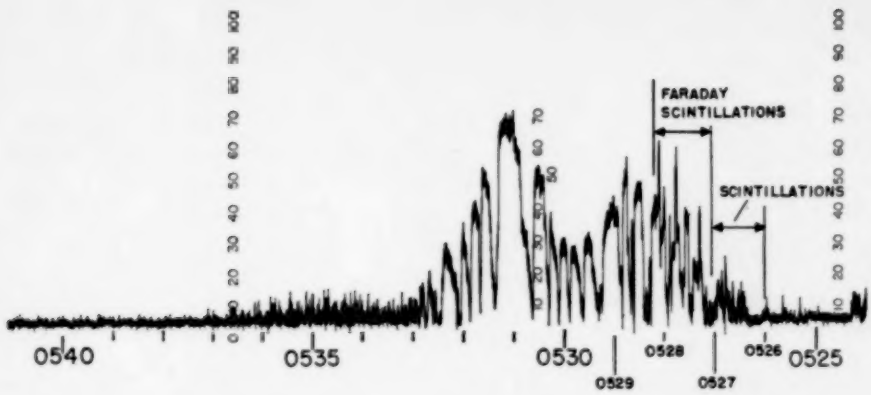


Fig. 15. Kiruna, Sweden, recordings of 1958 δ 2 revolution 7128, Sept. 30, 1959.

responsible, the Breisach 40 Mc/s signals would not have disappeared as they did. One explanation is that dropouts occur because passage of the satellite through an atmospheric region of high electron flux triggers the satellite transmitter, reduces its power, and thereby cuts the received signal to a low level.

Vernov *et al.*[2] report that Sputnik III was instrumented with a scintillation counter for photon recording. The intensity of the ionization was measured by varying the width of the transmitted pulse. From the Vernov paper it would appear that the average power in the original instrumentation would triple from the smallest pulse to the largest. The Kjeller fast-run recordings during the dropout period indicate transmitter modulation that may well have had an adverse effect on the power output. At this point in the satellite's rotations, the batteries may have been too weak to sustain the peak level required by a high ionization. Patches of the atmosphere evidently have a high electron flux, with electron energy to the value of 100 keV.

From observations at one station alone, it is difficult to determine whether the decrease being observed is due to the ionosphere or to a drop in transmitted power. The data obtained by Whitney *et al.*[3] shows that the 40 Mc/s signal at times persists during an ab-

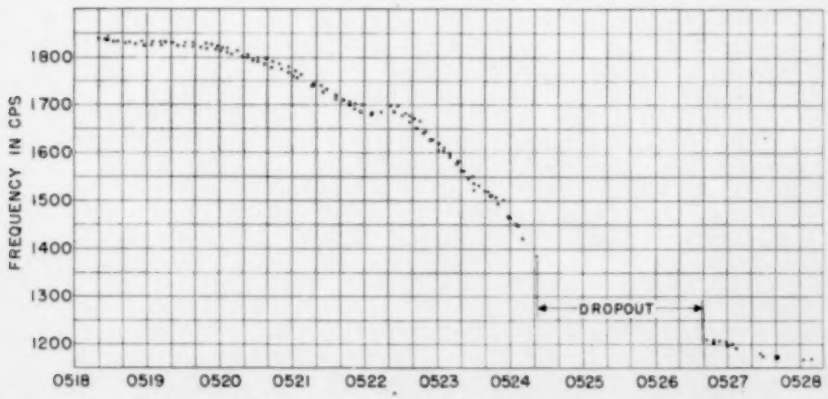


Fig. 16. Val-Joyeux, France, Doppler data on 1958 δ2 revolution 7128.

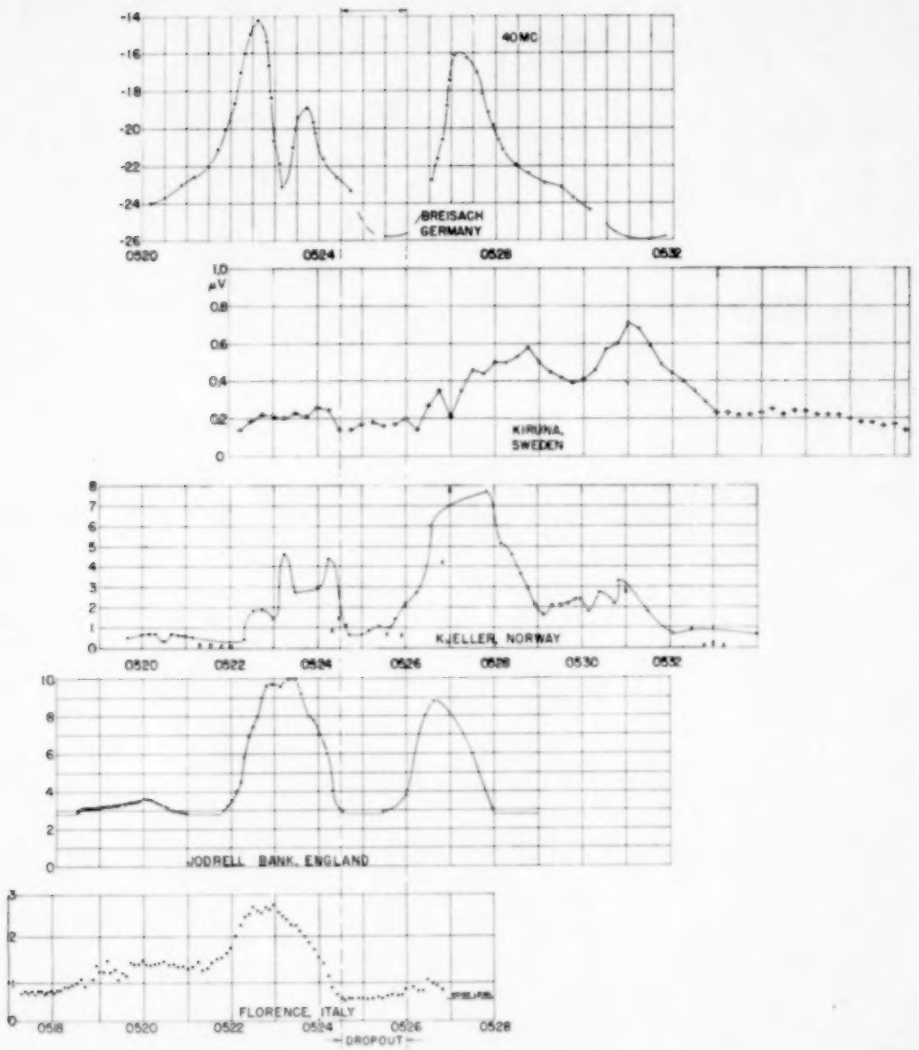


Fig. 17. Satellite 1958 $\delta 2$, revolution 7128, Sept. 30, 1959.

sorbing or scattering period, when all appearances indicated dropout.

A third factor that contributes to the difficulty of interpreting data or statistics solely from one station is the antenna pattern of the station itself. The Florence dipole antenna was oriented north-south and so its pattern contained a null to the north. During revolutions 7128 and 7172, the Florence signals were at a minimum just before the dropouts. Because of this decrease in amplitude just as the satellite entered a dropout period, the Florence observers may not have been able to determine whether a dropout had even occurred without other data from additional antennas, or at the harmonics of the signal, or from other observatories.

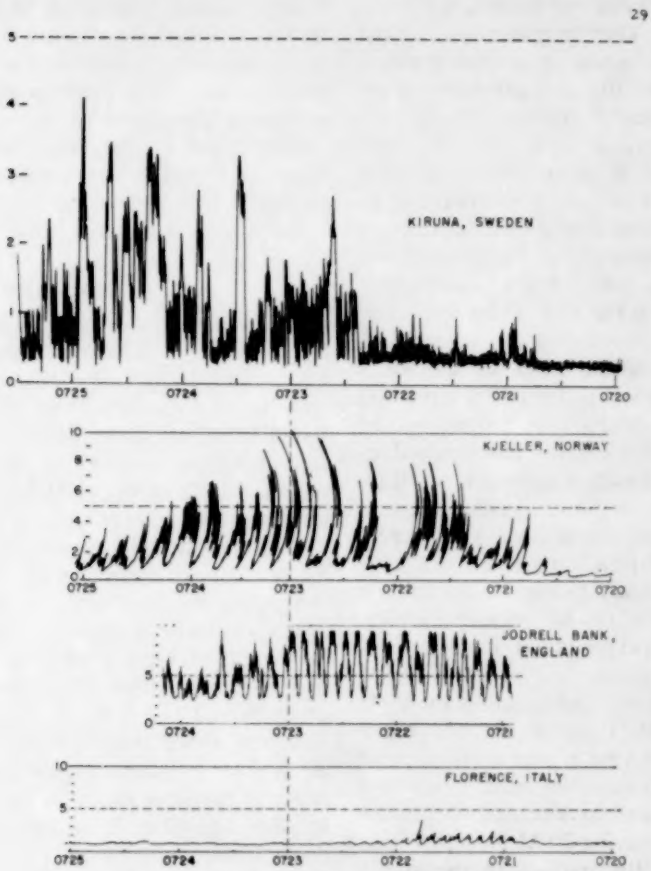


Fig. 18. The sharp signal decrease in the Kjeller recordings is evidence of a localized region of high ionization at 0722.

4.2. *The Scintillation region*

The ionospheric region producing scintillations cannot be defined too precisely. There is a slight elongation along north–south lines. During neither of revolutions 7172 and 7128 did Florence (43° 48′ 11″ N) observe scintillations; the other stations did. The scintillation region is not associated with the high-electron flux region that produces dropout since Faraday rotations were observed during dropout. It may however be that after a period of activity the high energy flux produces the scintillation "slabs." The extent of the subsatellite regions are 400 to 600 km along the north–south line, and somewhat less (perhaps 300 to 500 km) along the east–west line. Within this region there are further discontinuities that are less than 1 km in extent. The geometry indicates that the height may extend from a lower altitude ranging from 250 to 400 km to an upper altitude of 1000 km.

4.3. *The Scintillation mechanism*

Scintillation fluctuations observed in the 40 Mc/s signals of 1957 α (Sputnik I) similar to those observed on Sputnik III records have been described by Kent[4], who concluded that they are caused by the same irregularities in

the ionosphere as those responsible for radio star scintillations. The irregularities, about 1 km in size, occur mostly at heights between 270 and 320 km, usually concentrated toward the northern latitudes of Europe above 50°. Elongated along parallels of latitude, they tend to group themselves in areas several hundred kilometers in size. From the preliminary work of Frihagen and Tröim[5] at 20 Mc/s, there is evidence that the irregularities are considerably elongated along the lines of the geomagnetic field and can occur in the *E* or *F* layers.

There is some evidence that at least part of the amplitude fluctuation seen in 20 Mc/s linearly polarized signals is due to "polarization scintillation," a mechanism not apparent in radio star scintillations. If the amplitude fluctuations usually disclosed in radio star scintillation studies were the only ones present, it could be expected that they would always be superposed on the Faraday fading pattern and that the nulls of the Faraday fading would still be evident. There are, of course, records showing such scintillation superposed on Faraday fading, but the scintillation has in many cases obliterated the Faraday fading altogether. If irregularities of about 1 km in size can effect variations of 1–2 per cent in the total electron density along the propagation path, it may be enough to cause random fluctuations in the polarization of the arriving 20 Mc/s signal, and thus cause random fluctuations in the signal amplitude received by a linearly polarized (or partially linearly polarized) antenna. If the polarization of the arriving signal were to change more than about half a rotation, the Faraday fading pattern could be completely obliterated. When irregularities appear to be of the same size viewed from all directions, the polarization scintillation caused by a variation in total electron density will be greatest when the propagation path is along the magnetic field, and will vary as the cosine of the angle between path and field. When the irregularities are elongated parallel to the field, the amount of polarization scintillation will also depend on how the propagation path aligns with the irregularities.

Polarization scintillation will also take place if the ordinary and extraordinary components of the arriving wave travel by different paths and experience uncorrelated phase scintillations in doing so. This mechanism would be most serious at low elevation angles (and along the magnetic field), where the difference in the refraction of the two components at 20 Mc/s might be sufficient to produce different phase scintillations in the two paths.

5. SUMMARY

Amplitude and doppler recordings of signals from 1958 δ2 disclose several phenomena. From knowledge of the transmitter characteristics and observations of the transmitter modulation, it is suggested that dropout is caused by passage of the satellite through a region of high-energy particles. Rather sharply localized ionized regions have been recorded.

Scintillation regions have also been observed These "slabs" show a slight tendency toward elongation in the north–south direction. As nearly as can be determined from the geometry, the height of the regions may range from 250 to 400 km as a lower limit to 1000 km as the upper boundary.

A part of the amplitude fluctuation may well be due to polarization scintillation. Another cause of the scintillation is undoubtedly related to the scintillation of randomly polarized discrete sources.

Acknowledgements—This program was sponsored by the Electronics Research Directorate (AFRD) and Project Spacetrack (AFCCDD) through the European Office of the Air Research and Development Command. The analysis was co-ordinated by J. Aarons and H. E. Whitney of the Radio Astronomy Section of the Propagation Sciences Laboratory.

REFERENCES

1. A. Nicholl, private communication (1960).
2. S. N. Vernov, A. E. Chudakov, E. V. Gorchakov, J. L. Logachev and P. V. Vakulov, *J. Planet. Space Sci.* **1**, 86 (1959).
3. H. E. Whitney, H. Strick J. Aarons and J. Mott. *J. Geophys. Res.* **65**, 4210 (1960).
4. G. S. Kent, *J. Atmos. Terr. Phys.* **16**, 10 (1959).
5. J. Frihagan and J. Tröim, *J. Atmos. Terr. Phys.* **18**, 75 (1960).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 185–195. Printed in Great Britain.

THE INFLUENCE OF RESONANCE ABSORPTION IN THE SODIUM LAYER ON TEMPERATURE AND DENSITY MEASUREMENTS IN THE UPPER ATMOSPHERE

T. M. DONAHUE*
The University of Pittsburgh, Pittsburgh, Pennsylvania

J. E. BLAMONT and M. L. LORY
Service d'Aéronomie du C.N.R.S. Observatoire de Meudon, France

(*Received* 28 *September* 1960)

Abstract—Resonance absorption of sunlight by the earth's own sodium layer seriously affects line width and intensity measurements on an artificial sodium cloud ejected from a rocket. For an actual case where the altitude of the earth's sodium and its abundance were also measured it is shown that although the twilight measurement revealed an anomalously low and thin layer this is just what is needed to explain quantitatively the line shape and D_2/D_1 measurements obtained from the artificial sodium cloud. These results confirm the reliability of the twilight measurements. They also show that temperature and density measurements are possible by use of these techniques.

The measurements by Bricard and Kastler in 1944 giving the temperature of the earth's sodium are corrected for this effect. The result is a temperature of 200° at 80 km of altitude.

1. INTRODUCTION

The ejection of a trail of alkali vapor from a rocket fired during twilight permits in principle the observation and measurement of the wind velocity, the spectrum of turbulence, the atmospheric density (through a measurement of the diffusion coefficient), the rate of disappearance of the ejected atoms and the temperature of the atmosphere above 85 km. The temperature may be determined from a measurement of the width of the scattered resonance line provided the line has a Doppler shape. In practice the line is often not Gaussian for several reasons, one of which is that the optical thickness ($\tau_0 \equiv k_0 l$) of the cloud is too large. When τ_0 is smaller than unity the inaccuracy of the temperature measurement is not expected to be large especially since it is possible to correct the results for the broadening produced by multiple scattering. The procedure is then to measure independently and simultaneously the line shape and the optical thickness and to use the line shape for temperature measurements only when $\tau_0 < 1$. The optical thickness may be obtained from the intensity ratio of the fine structure components, D_2/D_1, as measured by means of a Fabry–Perot interferometer.[1] The variation of this ratio with time is a measure of the rate of disappearance of the free alkali atoms.[2] The line shape may be determined from an absorption measurement utilizing sodium vapor cells following a technique developed by Bricard and Kastler for the twilight flash.

The object of this paper is to point out that when the metal used is sodium one serious complication arises with regard to these two measurements because of the presence of the earth's own layer of sodium clouds through which the sunlight must pass to illuminate the artificial sodium (except for sufficiently high parts of the rocket trajectory and for sufficiently small angles of solar depression). This layer can be optically thick for the grazing path followed in twilight which can be as much as 70 times as long as the vertical path through the layer.[3–5] The spectrum of the sunlight incident on the cloud will be altered by absorption in this sodium layer and the intensity, line

* John Simon Guggenheim Fellow, 1960, Service d'Aéronomie, Observatoire de Meudon, Seine et Oise, France.

shape and doublet ratio of the light scattered by the cloud seriously affected.[3-9] It will be shown however that it is possible to predict these effects quantitatively and take them into account. To do so the altitude, distribution and concentration of the earth's sodium is determined with a magnetic scanning photometer[10] every time an artificial sodium cloud is formed. Examples of the appearance of these effects will be taken from the measurements made during the firing of a Veronique rocket at Hammaguir (Algeria), 2 March 1960, when the light scattered by the sodium released showed quite clearly the expected effects of this absorption. The agreement between the results predicted from the parameters of the sodium layer determined by the twilight flash observations and the observed properties of the artificial cloud provides an excellent check on the information gathered about the earth's sodium layer from ordinary twilight measurements.[9, 11, 12]

2. THE D_2/D_1 RATIO

The ratio in intensity of the two fine structure components in the light scattered by a thin sodium cloud would be 2 if the exciting spectrum were white. As the ratio of intensity at the bottom of the Fraunhofer lines is 0·85[5, 13] this ratio would become 1·7 but is seriously reduced in an optically thick cloud. It is a function of τ_0 decreasing toward unity as τ_b tends to infinity. This function has been computed for a simple cloud geometry and published elsewhere.[3] By virtue of this computation the observation of D_2/D_1 determines τ_0 for the cloud provided the cloud lies above the shadow of the earth's sodium. When the cloud is within this shadow absorption within the sodium layer reduces yet further the D_2/D_1 ratio. The geometry of the situation is illustrated in Fig. 1.

On the evening of March 2, the date of the launching discussed here, observations were made on the twilight flash. By an odd chance the characteristics of the layer that day were rather unusual. The amount of sodium was 2×10^9 atoms/cm^2, whereas the amount observed at that station over a period of several weeks before and after was ordinarily about $7{\cdot}5\times10^9$; the altitude was very low: 49 km in one direction, 70 km in another and 90 km in a third; the vertical extent was only 4 km.

For such a sodium layer: 4 km wide, containing 2×10^9 atoms/cm^2 and located in agreement with these observations very low—specifically at 62 km for reasons which will follow—the ratio τ/τ_v of the horizontal to the vertical optical thickness has been computed as a function of the height of passage above the earth.[3] This function, along with transmission functions for the D_2 and D_1 components computed by methods discussed in several of the papers

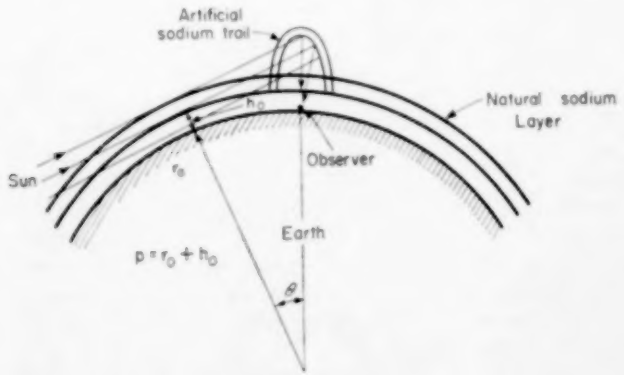


Fig. 1. Illumination of sodium rocket trail through the natural sodium layer.

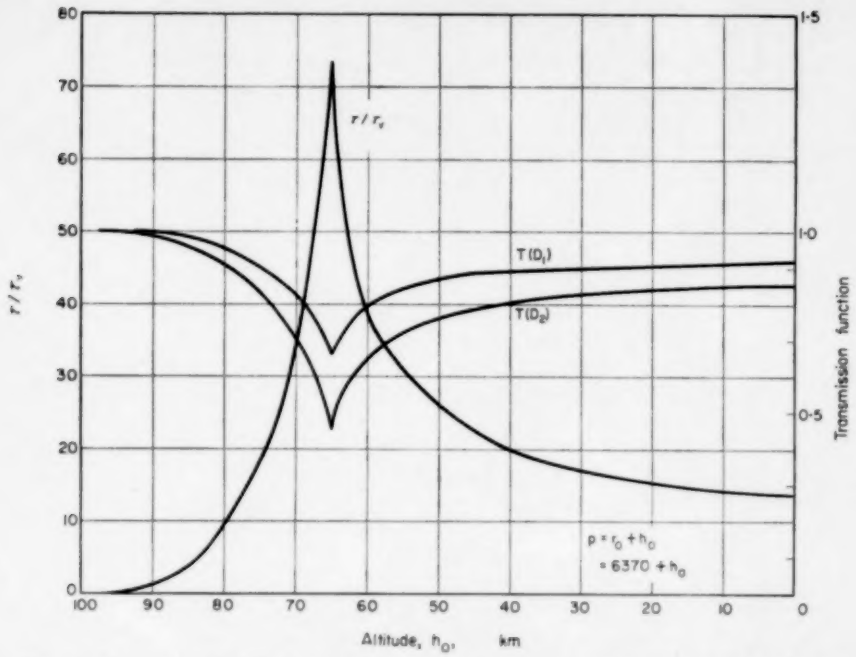


Fig. 2. Ratio of horizontal to vertical path entirely through a sodium layer 4 km wide at 62 km. Transmission functions for D_1 and D_2 lines when the layer contains 2×10^9 atoms/cm^2.

already referred to[4, 5, 9], is plotted in Fig. 2. All of these functions are very sensitive to a variation in the parameters of the layer as can be seen in Fig. 3. Since the height of the layer varies between 50 and 150 km, its width between 1 and 50 km and the amount of sodium from below 1×10^9 atoms/cm^2 to more than 25×10^9 circumstances are apt to change radically from one launching to another.

The rocket fired on 2 March 1960, rose to 188 km and ejected 280 gr of sodium (together with 1 kg of potassium). This small quantity was chosen in order to obtain an optically thin cloud of sodium. The cloud formed could be observed and studied with yellow filters. The D_2/D_1 ratio was measured at the levels 190, 147 and 110 km. The results are plotted in Fig. 4 (reproduced from Ref. 2). The ratio rises at first, then drops and finally rises again. This was just what was expected to happen as the result of the passage of the earth's sodium shadow across the cloud. From the curves of Fig. 2 the effect of the sodium shadow on D_2/D_1 may be computed as a function of the angle of solar depression for each of the three cloud levels. The result is shown in Fig. 4 along with the D_2/D_1 values corrected by this factor. After the correction the D_2/D_1 values rise smoothly toward the proper limiting value of 1·7. It would not be possible to shift the sodium layer more than 3 or 4 km or change its width and optical thickness appreciably without distorting the corrected D_2/D_1 curve seriously. It is for this reason that the exact height of 62 km was selected. The location of the minimum in the transmission function for the 190 km cloud level and a layer at 85 km is indicated in the figure.

Because of these effects lithium, whose abundance in the atmosphere is small, was released in an experiment on June 16, 1960. The variation of D_2/D_1 for lithium tends smoothly toward the limiting value of 2 in striking contrast to the behaviour of sodium.

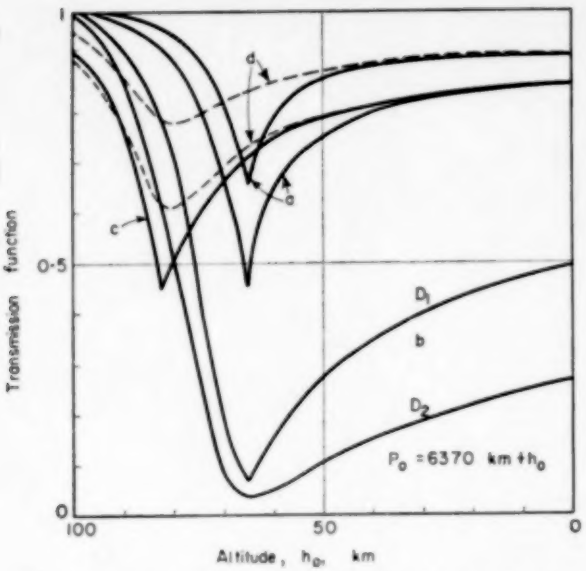


Fig. 3. Transmission functions for D_1 and D_2 (a) 4 km wide layer at 62 km containing 2×10^9 atoms/cm^2; (b) same containing 20×10^9 atoms/cm^2; (c) 4 km wide layer at 85 km containing 2×10^9 atoms/cm^2 (only D_2); (d) same 10 km wide.

3. THE LINE WIDTH

The temperature measurement is made by means of a modification of the Bricard–Kastler[14] absorption technique. This procedure was developed originally to measure the temperature of the natural sodium layer and to demonstrate that the D lines in twilight are excited by resonance scattering. Bricard and Kastler measured the absorption of the light from the twilight flash in a sodium vapor cell by means of a spectrograph. The technique in the present case is to photograph the sodium cloud in light which has passed through an absorption cell containing sodium vapor of known temperature, T_a and optical thickness τ_2 and also to photograph it simultaneously in light which has not passed through the absorbing sodium. If all spectral lines involved have a Doppler shape and if we adopt the notation

$$x^2=\frac{Mc^2}{2RT_c}\left(\frac{\nu-\nu_0}{\nu_0}\right)^2 \quad . \ . \ . \quad (1)$$

where M is the atomic weight of sodium, c the velocity of light, R the gas constant, T_c the temperature of the cloud and ν the frequency, then for a single hyperfine component of the D lines the light observed without the absorption cell is proportional to

$$I_1'=\int \exp-(x^2)\,dx \quad . \ . \ . \quad (2)$$

while that observed through the cell is—with the same constant of proportionality—proportional to

$$A_1'=\int \exp-(x^2)\exp-[\tau_2\exp-(b'x^2)]dx \quad . \ . \ . \quad (3)$$

where

$$b'=T_c/T_a \quad . \ . \ . \quad (4)$$

The ratio A'_1/I'_1 is then the ratio of the observed intensities and is a function of T_c only if b' and τ_2 are fixed. If A'_1 and I'_1 represent the absorbed and unabsorbed light for D_1 in equations (3) and (2), then A'_2 and I'_2 for

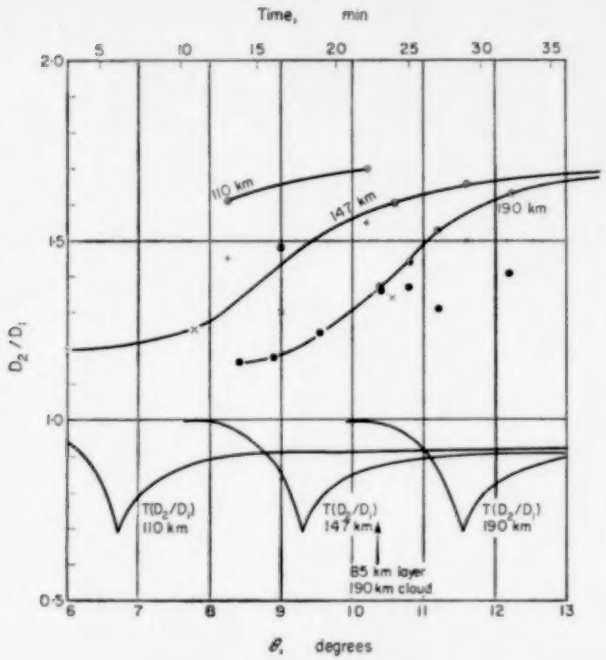


Fig. 4. The ratio of the transmission functions for D_2 and D_1 as a function of the angle of solar depression for light arriving at levels of 110, 147 and 190 km and layer conditions of Fig. 2. Experimental D_2/D_1 values: (●) 190 km (×) 147 km and (+) 110 km, and these values divided by the appropriate function taken from the lower part of the figure (⊙) 190 km, ⊗ 147 km and ⊕ 110 km.

D_2 are given by the same expressions with $2\tau_2$ replacing τ_2 in equation (3). As the pictures are taken in total light (D_2 and D_1) the absorbed absorption ratio called the factor of reduction becomes

$$R' = \frac{A'_1 + 1{\cdot}7A'_2}{I'_1 + 1{\cdot}7\ I'_2} \quad . \ . \ . \quad (5)$$

when the factor 1·7 again takes into account the fact that for an optically thin cloud twice as much light is scattered in the D_2 component but there is only 0·85 times as much coming from the sun at the bottom of the D_2 Fraunhofer line.

In practice the combined light of all hyperfine components is measured and the computation of the temperature from the ratio of absorbed to unabsorbed light obscured becomes tedious. Bricard and Kastler make use of the empirical result of Zehden[18] according to which the absorption occurs as though the hyperfine structure of each fine structure component were not resolved and those line were Gaussian with effective cross sections of

$$\sigma_{o1} = 0{\cdot}40 \times 10^{11}\ \text{cm}^2 \text{ at the center of } D_1$$

$$\sigma_{o2} = 0{\cdot}80 \times 10^{11}\ \text{cm}^2 \text{ at the center of } D_2$$

This analysis evidently assumes that all spectral lines are Gaussian. If the optical thickness of the cloud, τ_o, is not negligible it is possible to take into account the broadening produced by multiple scattering (by methods which will be published separately) if τ_o is known, for instance from the D_2/D_1 measurements.

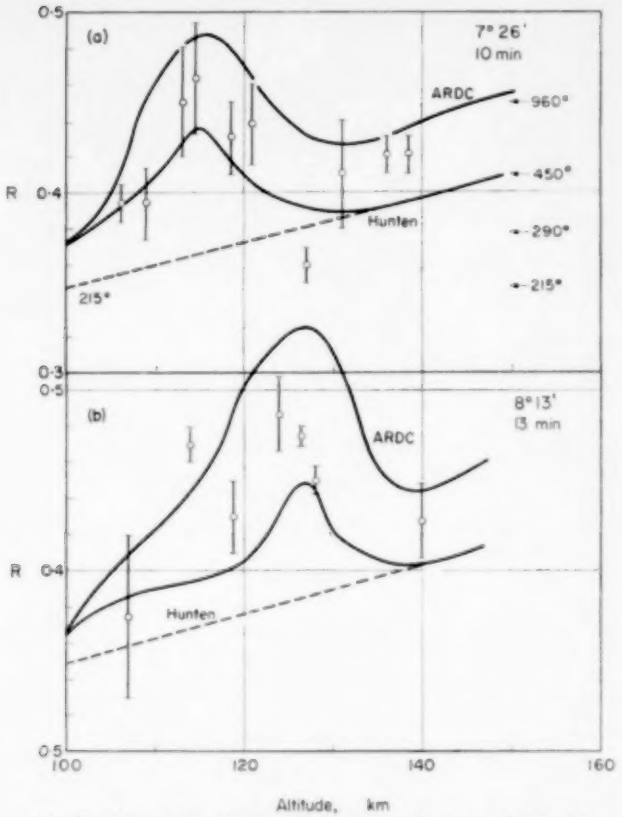


Fig. 5. Observed reduction factors as a function of altitude for two angles of solar depression. The theoretical curves are shown with (—) and without (....) the effects of absorption in the earth's sodium layer. Temperature scale is for the latter case only. The curve labelled Hunten is for the temperatures deduced by Hunten[16] from various spectroscopic observations on the airglow and aurorae. The upper curve is what would be predicted from the ARDC[17] model atmosphere.

The reduction factor R was obtained 5 times during the 30 min after firing of the rocket already discussed at altitudes between 100 and 150 km. The results are plotted in Figs. 5–7. R shows a peak at 117 km when the sun is 7°26′ below the horizon but gets progressively higher as the sun sets. This again is explicable as an effect of the absorption in the earth's sodium layer. Effectively the result is to broaden (even reverse) the resonance line. The observed maximum locates the place where the deepest sodium shadow is cast on the cloud. That shadow then rises as the sun sets.

When the sunlight must pass through a part of the natural sodium whose effective[15] optical thickness is τ_1 at the center of the D_1 component to reach the portion of the cloud under observation the expression (2) for I'_1 becomes

$$I_1 = \int \exp - [\tau_1 \exp - (x^2)] \exp - (ax^2)\, dx \quad . \quad (6)$$

Here

$$a = T_e / T_c \quad . \quad . \quad . \quad (7)$$

where T_e is the temperature of the layer and x^2

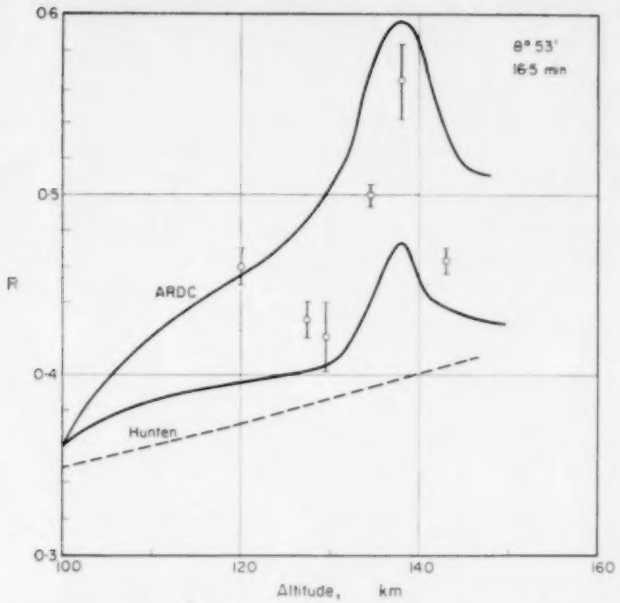


Fig. 6. Experimental and theoretical reduction factors 16·5 min after firing. The maximum at 120 km here and at 114 km in Fig. 5 (b) would imply a second cloud at 47 km. It is interesting that the twilight flash also located sodium at 49 km in one direction. However, at the 114-km level there is a sharp bend in the trail caused by the wind and this maximum may result from the great optical thickness of the trail.

is now defined in terms of T_e rather than T_c:

$$x^2 = \frac{Mc^2}{2RT_e}\left(\frac{\nu-\nu_o^2}{\nu_o}\right) \quad . \quad . \quad . \quad (8)$$

Similarly the expression (3) for A'_1 with absorption in the cell taken into account becomes:

$$A_1 = \int \begin{matrix} \exp-[\tau_1 \exp-(x^2)] \exp-(ax^2) \\ \exp-[\tau_2 \exp-(bx^2)]\, dx \end{matrix} \quad . \quad . \quad (9)$$

where

$$b = T_e/T_a \quad . \quad . \quad . \quad (10)$$

I_2 and A_2 may be computed also for the D_2 line when τ_2 and τ_1 are replaced by $2\tau_2$ and $2\tau_1$ and hence the new reduction factor

$$R = \frac{A_1 + 1{\cdot}7\ A_2}{I_1 + 1{\cdot}7\ I_2} \quad . \quad . \quad . \quad (11)$$

This has been done and R plotted as function of τ_1 for several values of a and b (Fig. 8).

If the natural sodium layer is taken again to lie at 62 km and to contain the same amount of sodium as before and τ_1 read off Fig. 2 as a function of h_o, that is as a function of the altitude of the artificial sodium cloud for each angle of solar depression, R may be obtained and plotted against altitude as in Figs. 5–7. The value of R observed are also plotted in these figures. Each point recorded represents an average over three to six observations spread over 2 km. The correction for multiple scattering has not been applied to these observations.

In all of the computations b was taken to be 0·6, i.e. $T_A = 480°$ and $T_e = 288°$. T_e was then permitted to vary from 215° at 100 km to 450° at 147 km. The temperature of the layer was chosen to be that of the conventional temperature at 62 km. If, however, it were chosen to be 190° there would be little significant change in

the results considering the uncertainties in the measurements and in the correction for multiple scattering.

There can be little doubt that the variation observed in R is caused by the natural sodium. Also it is clear that the twilight sodium observations have located and defined the terrestrial sodium layer with a high degree of reliability. To add force to this conclusion there is the happy circumstance that the twilight observation detected a very thin layer lying at extremely

4. TEMPERATURE MEASUREMENTS OF BRICARD AND KASTLER

These considerations also apply to the measurement of the temperature of the natural sodium by Bricard and Kastler[14]. In their excellent and very complete paper on the subject these authors consider the possible effect of absorption by the earth's sodium on the intensity and D_2/D_1 ratio in the twilight flash but conclude that it would not be serious if there were

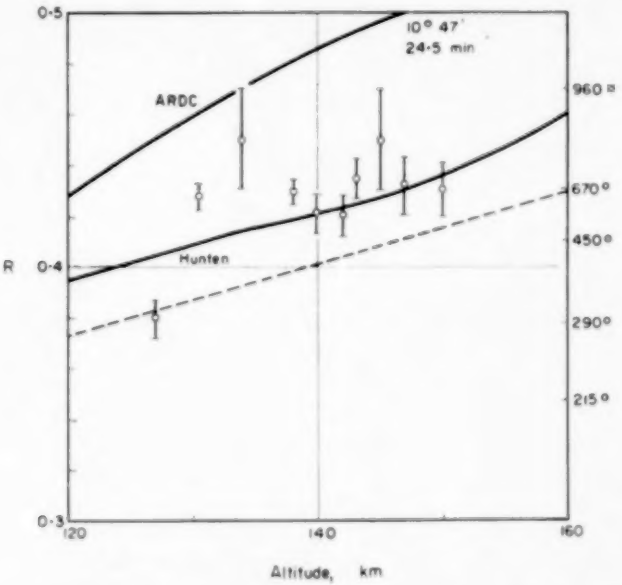


Fig. 7. Experimental and theoretical ratios late in twilight.

low altitude and containing very little sodium. All of these properties are quantitatively required to explain the variation of D_2/D_1 and R in the artificial cloud. This fact supports strongly the reliability of the method of height determination and the calibration of Blamont's Zeeman photometer which has been employed by Blamont and Donahue [8, 9, 11]. It is also direct evidence of the great departures from 85 km for the sodium layer that they have reported.[11]

less than 10^{11} atoms/cm^2. This is true for those effects which depend on transport vertically through the layer. Subsequent study has shown that for the great optical paths encountered in twilight the influence of resonance absorption can be important when there are as few as 2×10^9 atoms/cm^2. This is the case in particular for the measurement of temperature. Instead of the factors appearing in equations (6) and (9) the ratio R is determined approximately by factors such as

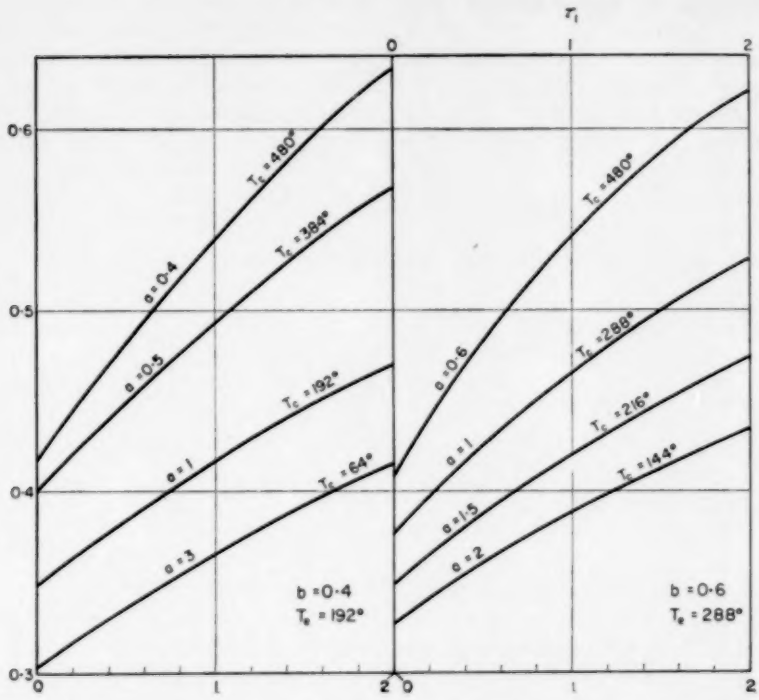


Fig. 8. Calculated values of R for various cloud and layer temperatures as a function of the optical thickness of the layer traversed.

$$A_1 = \int \exp-[\tau_1 \exp-(x^2)] \exp-(x^2) \exp-[\tau_2 \exp-(bx^2)]\, dx \quad . \quad (10)$$

$$I_1 = \int \exp-[\tau_1 \exp-(x^2)] \exp-(x^2)\, dx \quad . \quad (11)$$

where τ_1 in the optical thickness from the sun to the effective center of the sodium layer at the point of observation. Rigorously these quantities ought to be computed for each hyperfine component by the methods of Donahue[5] or Chamberlain *et al.*[8] and the absorption function should be the curve given for the appropriate temperature by Chamberlain *et al.*[8] However the error committed by using effective cross sections while neglecting the hyperfine structure[14] and in selecting an effective center of the layer instead of integrating along the direction of observation is not serious except for very thick layers.

Again:

$$R = \frac{A_1 + 1{\cdot}7\, A_2}{I_1 + 1{\cdot}7\, I_2} \quad . \quad . \quad . \quad (12)$$

This formula uses the factor 1·7 in place of the factor 2 which Bricard and Kastler used for lack of knowledge concerning the intensities at the bottom of the Fraunhofer lines. Fortunately the error committed by use of 2 in place of 1·7 partly compensates for the error committed by neglecting the broadening caused by absorption. Use of the old formula gives a temperature too low when there is less than about 4×10^9 atoms/cm^2 and too high for abundances higher than that.

Bricard and Kastler obtained the following results:

Jan. 19, 1944 $\tau_2 = 0{\cdot}57$, $T_e = 480°$K, $R = 0{\cdot}500$

Jan. 30, 1944 $\tau_2 = 0{\cdot}77$, $T_e = 480°$K, $R = 0{\cdot}403$

They concluded that the temperature T_e was $240 \pm 50°K$.

We have computed the values of R given in Table 1, taking the effective value of τ_1 to be $13 \times \tau_V$ which it will have at 6·5° solar depression for almost any model of the sodium layer. Several values of L_0, the sodium abundance, are assumed.

Table 1.

τ_2	τ_1	L_0 (atoms/cm²)	b	T_e	R	without absorp.	with factor 2
0·77	0·2	$3{\cdot}8 \times 10^9$	0·4	192°	0·385	0·372	0·350
0·77	0·2	$3{\cdot}8 \times 10^9$	0·5	240°	0·401		0·390
0·57	0·2	$3{\cdot}8 \times 10^9$	0·5	240°	0·496		
0·77	0·4	$7{\cdot}7 \times 10^9$	0·4	192°	0·401		
0·77	0	0	0·67	320°	0·400	0·400	
0·77	1·0	19×10^9	0·25	120°	0·400		

Experience with the seasonal variation at the latitude of the Pic du Midi indicates that toward the end of January the abundance should be part of the way down to its February minimum.[13] It should have been somewhere between 4×10^9 and 10×10^9 atoms/cm². This should give a temperature between 235° and 170°. The most likely value—corresponding to about 7×10^9 atoms/cm²—is 198°. The altitude at this time of the year was probably close to 80 km.

This experiment is one which would well bear repeating if it should be accompanied by abundance and altitude measurements. This is particularly true now that it is realized over what a wide range the sodium altitude can vary in some latitudes.

5. CONCLUSIONS

(1) The effect of resonance absorption in the earth's sodium layer on the D_2/D_1 ratio and line width of an artificial cloud of sodium are easily detectable. They can be compensated for if the location and abundance of the earth's sodium are measured. A better procedure might be to use lithium or potassium for the measurements: lithium for the D_2/D_1 measurements (since one component of the potassium doublet is erased by water vapor) and potassium for the line width (since lithium attacks glass). These two elements should be released from the same rocket or at least at the same place during the same twilight. An alternative would be to try to use an interferometer to obtain the lithium line shape.

(2) The method used by Blamont and Donahue to determine the sodium altitude and abundance is reliable. In particular the earth's sodium is to be found at times as low as 60 km—and higher than 150 km. The abundance in the latitude of 40° rises above 25×10^9 atoms/cm² sometimes.

(3) The temperature of the earth's sodium may be measured by the method of Kastler and Bricard provided the effect of absorption by the telluric sodium is accounted for. Their results corrected give a temperature of about 200° at about 80 km. Their method combined with height and abundance measurements is capable of measuring the atmospheric temperature from 50 to 150 km, account being taken of the great variability in altitude of the sodium during the year.

(4) The temperature between 80 and 150 km measured by the methods discussed in this paper is certainly in disagreement with that of the ARDC model atmosphere[16], but is not in disagreement with the concensus of spectroscopic values arrived at by Hunten[17]. In particular, the new value deduced from the Bricard and Kastler measurement at 80 km lies on the curve drawn by Hunten. It is clear, particularly from Fig. 7 where the corrections for multiple scattering ought to be small, that

Table 2.

Altitude in km	Temperature in °K
80	198°K
100	215° ± 25°
110	240° ± 30°
120	275° ± 45°
130	325° +65°, −50°
140	400° +80°, −70°
150	515° +150°, −115°

the predictions of R from the temperature curve proposed by Hunten are in reasonable agreement with these observations, rough as they are. The temperatures we obtain are summarized in Table 2. The temperature variation is like that which Nicolet obtains for one of his models[18], the one which is based on dissociation of O_2 above 100 km, diffusion from 115 km and appreciable heat conduction from above.

REFERENCES

1. J. E. Blamont, *C.R. Acad. Sci., Paris* **250,** 567 (1960).
2. J. E. Blamont, M. L. Lory and G. Courtes, *Ann. Geophys.*, **16,** 435 (1960).
3. T. M. Donahue and R. Resnick, *Phys. Rev.* **98,** 1622 (1955).
4. T. M. Donahue, R. Resnick and V. R. Stull, *Phys. Rev.* **104,** 873 (1956).
5. T. M. Donahue and V. R. Stull, *Ann. Geophys.* **15,** 481 (1959).
6. G. I. Galperin, *Astron. Zh.* **33,** 173 (1956).
7. J. W. Chamberlain, *J. Atmos. Terr. Phys.* **9,** 73 (1956).
8. J. W. Chamberlain, D. M. Hunten and J. E. Mack, *J. Atmos. Terr. Phys.* **12,** 153 (1958).
9. J. E. Blamont, T. M. Donahue and V. R. Stull, cited in *Ann. Géophys.* **14,** 282 (1958).
10. J. E. Blamont, *The Airglow and the Aurorae* (edited by E. B. Armstrong and A. Dalgarno), p. 99, Pergamon Press, London (1956).
11. T. M. Donahue and J. E. Blamont, *Proc. Symp. on Aeronomy, Copenhagen 1960*, to be published in *Ann. Geophys.*
12. J. E. Blamont, T. M. Donahue and W. Weber, *Ann. Géophys.* **14,** 282 (1958).
13. J. A. Scrimger and D. M. Hunten, *Canad. J. Phys.* **35,** 918 (1957).
14. J. Bricard and A. Kastler, *Ann. Geophys.* **1,** 53 (1944).
15. W. Zehden, *Z. Phys.* **86,** 555 (1933).
16. R. A. Minzer, K. S. W. Champion and H. L. Pond, Air Force Surveys in Geophysics No. 115, Geophysics Research Directorate, Bedford, Massachusetts, 1959 (unpublished).
17. D. M. Hunten, *Symposium on Aeronomy at Copenhagen*, 1960 (to be published).
18. M. Nicolet in *Physics of the Upper Atmosphere* (Ed. by J. A. Ratcliffe), p. 17, Academic Press, Inc., New York and London.

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 196–201. Printed in Great Britain.

COMMENTS ON TIDAL WINDS IN THE HIGH ATMOSPHERE

B. HAURWITZ

Department of Astro-Geophysics, University of Colorado, Boulder, Colorado

(*Received* 2 *September* 1960)

Abstract—Radio meteor observations have shown the existence of strong semidiurnal tidal winds in the high atmosphere. The accuracy of published tidal-wind determinations is discussed on the basis of computations of the radius of the probable-error circle. It is shown that the strong variations of the tidal winds with the seasons and with elevation may be regarded as real, although their actual magnitude cannot yet be regarded as known reliably. The implications of these variations for the solar air tide are discussed.

1. TIDAL WINDS IN THE LOWER ATMOSPHERE

In conjunction with the solar semidiurnal pressure oscillation a wind variation of the same period is known to occur. At the earth's surface these periodic winds are weak, with amplitudes of less than 1 m/sec, so that they can be found only by a statistical analysis of a long series of data. Perhaps the best and least unknown determination* is that by Chapman[1] for the solar semidiurnal wind at Mauritius (20·1°S, 57·5°E) from about 16 years of data; this he obtained in the course of his (so far unique) determination of the wind system of the lunar air tide:

$$u = 13{\cdot}7 \text{ cm/sec} \sin (30t + 86°)$$

$$v = 28{\cdot}0 \text{ cm/sec} \sin (30t + 271°)\dagger$$

Here, and in the following, u will mean a wind *to* the south (a north wind in meteorological terminology), and v a wind *to* the east (a west wind in meteorological terminology). The local mean time in hours is denoted by t, and the oscillations will always be expressed in the above form, namely,

$$A \cdot \sin (30t + \alpha)$$

with A for the amplitude and α for the phase angle. A measure of the statistical reliability and accuracy of the expression for the oscillation is given by the radius of the probable-error circle, which is 0·6 cm/sec in the case of the Mauritius winds. A determination of such an oscillation may be considered fairly satisfactory when the amplitude is at least 3 times this radius.

To determine the radius of the probable-error circle the oscillation must be written in the form

$$a \cos 30t + b \sin 30t$$

(with $a = A \sin \alpha$, $b = A \cos \alpha$, by comparison with the foregoing expression). Let a_i, b_i be the harmonic coefficients of the n individual determinations; $\bar{a}$, $\bar{b}$ their arithmetic means, and compute $\Delta a_i = a_i - \bar{a}$, $\Delta b_i = b_i - \bar{b}$. Then the radius of the probable-error circle for the individual determination is

$$r^* = 0{\cdot}83 \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[(\Delta a_i)^2 + (\Delta a_i)^2 \right] \right\}^{\frac{1}{2}},$$

and the radius of the probable-error circle for the mean

$$r = r^* / n^{\frac{1}{2}}$$

The value of r decreases as the number of observations increases. If the a_i, b_i are plotted on a "harmonic dial" and if a circle with radius r^* is drawn around the centroid $\bar{a}$, $\bar{b}$, half the points lie inside the circle, the other half outside, provided that the number of determinations is large enough so that the statistical laws apply. In the cases discussed below the number of data is marginal, but the statistical rule for r^* holds nevertheless pretty well.

* Prof. Chapman has called to my attention that another determination of the solar semidiurnal wind has been made for St. Helena (*Rep. Brit. Met. Comm.*, 1908).

† In the reference the phase angle is given as 91°, but inspection of the Mauritius data shows that this value refers to the component *from* the east.

For the lower stratosphere Johnson[2] has made a determination of the semidiurnal winds at the 150 mb level (approximately 15 km) over England showing winds of the same order of magnitude as at Mauritius, but no measure of the accuracy of the determination is available. Other determinations of semidiurnal wind periods in the lower atmosphere are unsatisfactory, either because of the short duration of the series of data or because the wind data are not analysed according to components.

2. TIDAL WINDS IN THE HIGH ATMOSPHERE

Sufficient wind data to demonstrate the existence of the semidiurnal period of the wind are lacking completely until an altitude of 75–80 km is reached. At these levels and above, much more frequent wind measurements have been made for some years by radio methods, especially by observing reflections from ionized meteor trials which drift with the ambient atmosphere These observations have shown that the semi-diurnal oscillation at these altitudes is much stronger than near the ground and can be recognized from a few days' observations, contrary to conditions at the surface of the earth. This increase of the amplitude of the tidal wind period with elevation is what one would expect from the theory of atmospheric oscillations.

Nevertheless considerable aperiodic changes must also be superimposed on the tidal wind variations in the high atmosphere, and the true tidal variations can only be ascertained if a sufficiently long series of observations is available to permit elimination of the aperiodic changes and computation of the probable-error circle, or an equivalent statistical measure to express the accuracy of the determination. This accuracy has nothing do with the method of wind observation, but rather with the effect of the other influences, the background noise overlying the periodic wind variations.

The largest number of data to determine tidal wind variations at high levels has been obtained by means of reflections from radio meteor trails for Jodrell Bank near Manchester, England, by Greenhow and Neufeld[3] and for Adelaide, Australia, by Elford[4]. These observations refer to heights roughly between 80 and 100 km. They are sufficient to permit at least a first estimate of the radius of the probable-error circle. Attention will therefore be given largely to these observations.

It is of interest to demonstrate in a quantitative fashion how much more pronounced the semi-diurnal (and the diurnal) wind variation is at greater heights than at the ground. Let A_u and A_v denote the amplitudes of the semidiurnal wind variations in the directions towards south and east, and u_m and v_m the corresponding mean winds. Then, for the twenty-four individual published cases from Jodrell Bank and the eighteen individual cases from Adelaide,

$A_u > u_m$ in 35 cases $\quad A_v > v_m$ in 22 cases

$A_u < u_m$ in 7 cases $\quad A_v < v_m$ in 20 cases

The preponderance of the semidiurnal terms over the mean values is, of course, particularly impressive for the meridional wind direction (u), since the mean wind components in this direction are considerably weaker than the zonal wind components (v).

For Jodrell Bank twenty-four individual determinations of the semi-diurnal wind variation, each for a 2-day period, have been published. For Adelaide wind periods have been given for different height intervals, but only the layer from 85 to 94 km comprises enough material (eighteen cases) for a rough determination of the probable-error circle.

Since the semidiurnal air tide shows seasonal variations, the available dates are here subdivided into three seasons:

E (equinoctial) months:
September, October, March, April.

D months:
November, December, January, February.

J months:
May, June, July, August.

This division is customary whenever a more detailed subdivision of the material is impracticable because of the scarcity of data. The following Table 1 shows for both stations the

Table 1. Semidiurnal wind variations at high levels

Jodrell Bank, England; 53·2°N, 2·3°W; 80–100 km

		u, positive to S			*v*, positive to E		
	n	*A*, m/sec	α	*r*, m/sec	*A*, m/sec	α	*r*, m/sec
Total	24	15·2	106°	±3·2	9·1	204°	±2·4
E	10	11·8	152°	±5·0	7·6	237°	±3·3
D	6	31·1	79°	±4·0	22·4	181°	±3·5
J	8	14·8	107°	±2·3	4·2	223°	±3·0

Adelaide, Australia; 34·9°S, 138·6°E; 85–94 km

		u, positive to S			*v*, positive to E		
	n	*A*, m/sec	α	*r*, m/sec	*A*, m/sec	α	*r*, m/sec
Total	18	4·8	193°	±3·4	1·9	106°	±4·0
E	7	3·0	210°	±3·9	3·7	85°	±5·6
D	6	11·6	266°	±4·5	11·0	254°	±2·8
J	5	17·1	136°	±6·2	13·9	84°	±9·2

number of observations, n, the amplitude, A, the phase constant, α, and the radius, r, of the probable-error circle of the mean. For the individual values from which the means are computed r^* is $\sqrt{n}$ times larger.

3. STATISTICAL RELIABILITY OF HIGH-LEVEL TIDAL-WIND DETERMINATIONS

Since the number of cases is small the values of the amplitudes, phase angles, and radii of the probable-error circles are not very certain. But the radii given in Table 1 provide at least a rough quantitative estimate of the reliability of the present and similar determinations of high-level tidal wind variations. As stated before, a determination may be regarded as fairly satisfactory if $A \geqq 3\ r$. For Jodrell Bank the majority of amplitudes satisfy this rule, but for Adelaide only the eastward component during the *D* months is above the critical value. It is not implied by this comparison between A and r that there are no tidal variations in the high atmosphere, but merely that they are not yet well determined.

A rough estimate of the number n of cases required for a satisfactory determination can be made from Table 1 since the error under normal conditions should decrease as $n^{-\frac{1}{2}}$. For instance, for the eastward component at Adelaide during the *J* months $A/r = 1{\cdot}5$. Thus, if the number of individual determinations were increased fourfold the radius of the probable-error circle would be decreased by one half so that $A/r = 3$. Again, however, the qualification must be added that this estimate is not very reliable because of the small size of the sample.

The semidiurnal wind components for Jodrell Bank and Adelaide are represented graphically in Figs. 1 and 2 in the form of harmonic dials. The positions of the annual mean are indicated by crosses, for which the error circles are not plotted in order to keep the figures less complicated. The figures and Table 1 show that there is a general tendency for the amplitudes of the semidiurnal wind components to be largest during the winter, and smallest during the summer months; there are also pronounced variations of the phase angles. Most of these seasonal changes are large enough to be considered real despite the large probable errors. Unfortunately, no direct determinations of the seasonal variations of surface tidal winds have been made. In order to compare conditions at upper levels with those at the surface, it is, therefore, necessary to make use of the seasonal variation of the semidiurnal surface pressure

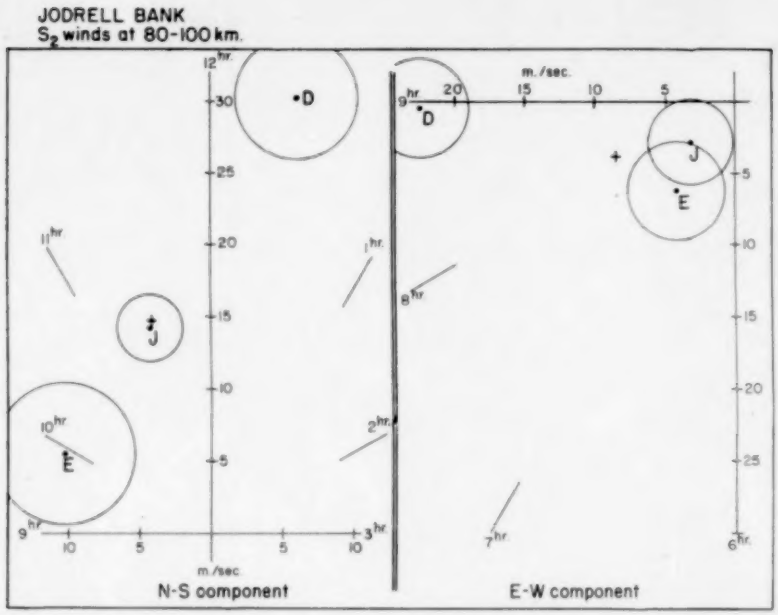


Fig. 1. Harmonic dial of solar semidiurnal winds at Jodrell Bank, England, height 80–100 km. (Here and in Fig. 2 the units of lengths are m/sec, and the hours corresponding to the time of maximum are indicated by straight lines along the margins of the figures). The error circles refer to the mean for the season.

oscillation.[5] The seasonal changes in the wind amplitudes which would be expected from these are in the same direction as the observed ones but much smaller. Thus, not only does the average semidiurnal variation appear to increase with elevation, but also its seasonal variation.

The radio meteor method permits discrimination between the different heights at which winds are measured. Accordingly both Elford[4] and Greenhow and Neufeld[6] have also determined the height variation of the semidiurnal wind components (as well as the mean and diurnal winds) for a layer reaching from 85 to 100 km. For Jodrell Bank the amplitude of both wind components increases with elevation at the rate of about 1·5 m/sec/km in winter, and approximately by a third of this value in summer. Similarly the phase angle increases, approximately at the rate of 7°/km (0·23 hr/km) in winter and 3°/km in summer. Elford's data for Adelaide give similar results, at least with respect to the amplitude increase and its seasonal variation.

For Jodrell Bank the height variations were found from a set of observations different from those by which the probable errors given in Table 1 were found. These newer observations were made between September 1954 and August 1955, when radio echoes were continuously recorded 2–3 days each month. Since individual values of the semidiurnal wind variations for the more recent observations are not published it is not possible to determine the radius of the probable-error circle. It can be assumed, however, that this radius is at least not larger than that for the individual values for the previous year, because during that year about 2 days' observations of echoes were used to find the wind oscillations, while during the later period 2–3 days were used. From Table 1 one finds that this radius is roughly 11 m/sec for the individual determination. Also, for the earlier period observations from rather different alti-

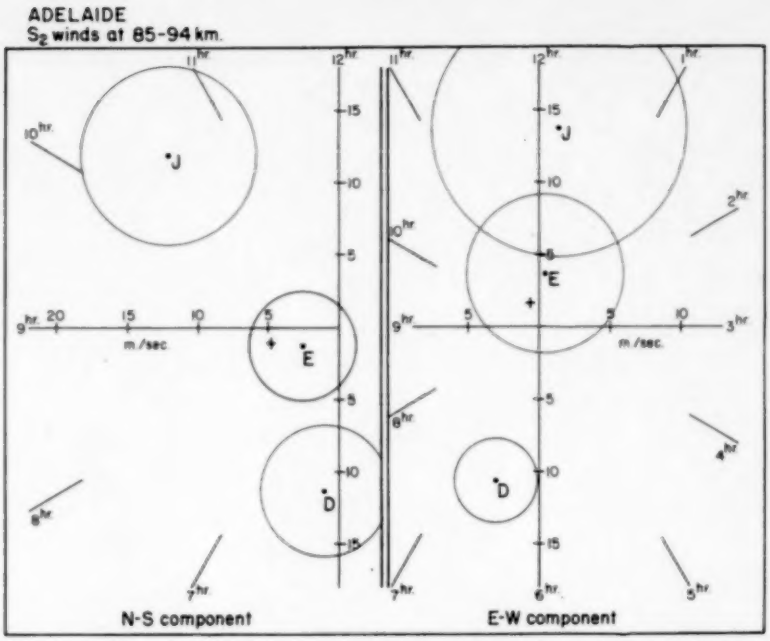


Fig. 2. Harmonic dial of solar semidiurnal winds at Adelaide, Australia, height 85–94 km.

tudes were combined, so that a rather inhomogeneous material was brought together. On the other hand, with the newer data, a separation according to layers was made, so that the material is probably more homogeneous. It can therefore be assumed that the height variations of the semidiurnal wind components as found by Greenhow and Neufeld and Jodrell Bank are real, at least as far as the general trend of the amplitude and phase angle changes are concerned. There is evidently a considerable amount of noise in the data as become apparent from an inspection of the hodographs in Ref.[6], especially in those cases where values for the same month during two different years are available. Consequently, a measure for the statistical reliability of the results is still highly desirable.

4. RELATION TO TIDAL THEORY

Such a measure, the radius of the probable-error circle, will be important for the further development of tidal theory. It has been pointed out by various writers that the increases of the amplitude and of the phase angle with elevation confirm the resonance theory of the solar semidiurnal tide, which requires a large magnification and hence quite accurate tuning between the period of the generating force, tidal and thermal action of the sun on the atmosphere, and a free period of the atmosphere. But the vertical temperature distribution of the atmosphere does not show the high temperature in the middle mesosphere required to produce such an accurate tuning. It is pertinent in this connection to refer to the explanation of the large solar semidiurnal oscillation given by Siebert[7]. According to this theory the large amplitude can be accounted for without large magnification, if one takes into account the fact that the atmosphere is not only warmed from the ground upward, but also by the absorption of direct solar radiation. In this manner not only the lowest few hundreds of meters but the atmosphere up to very great heights will be affected by the warming process. Hence, because of the larger magnitude of the generating force little resonance magnification is required

to produce the observed surface pressure change, and the tuning requirement is much less rigorous than in the older resonance theory.

According to theory [8] oscillations of a more realistic atmosphere than that required by the resonance theory may also show an increase of the wind speed and advance of the phase of the wind oscillation with elevation at higher levels and the phase reversal between the lowest and higher layers. It is therefore not possible to deduce from the general vertical distribution of the semidiurnal wind components, which of the theories of the solar semidiurnal oscillations is preferable. Such a choice would require much more detailed and accurate data than are currently available. These wind data should be distributed over a sufficient range of altitudes and latitudes. Because, especially at high altitudes, there seems to be a very pronounced seasonal variation of the semidiurnal wind oscillations, the data should further be sufficiently numerous to allow an appropriate grouping according to seasons.

Attention should also be given to the lunar tide since it is, at least from a theoretical viewpoint, easier to analyse, because here the forcing function, entirely due to gravity, is known exactly.

The amplitudes of the solar semidiurnal wind oscillations in the high atmosphere are of the same magnitude as the mean winds. Therefore, a reasonably complete knowledge of these oscillations is indispensable, not only from a theoretical viewpoint, but also for practical purposes whenever information on the atmospheric field of motion at high levels is required.

Acknowledgement—The research reported in this document has been sponsored by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command under contract AF 19(604)-5488.

REFERENCES

1. S. CHAPMAN, Presid. Adr., Assoc. of Meteorol., *U.G.G.I. 8th Gen. Assembly, Oslo, 1949.* Also in *Compendium of Meteorology* p. 510; American Meteorol. Soc., Boston (1951).
2. D. H. JOHNSON, *Quart. J. Roy. Met. Soc.* **81**, 1 (1955).
3. J. S. GREENHOW and E. L. NEUFELD, *Phil. Mag.* (7)**46**, 549 (1955).
4. W. G. ELFORD, *Planet. Space Sci.* **1**, 94 (1959).
5. B. HAURWITZ, and G. M. SEPULVEDA, *J. Met. Soc. Japan* 75th Annivers. Vol., 149 (1957).
6. J. S. GREENHOW and E. L. NEUFELD, *Phil. Mag.* (8)**1**, 1157 (1956).
7. M. SIEBERT, *Ber. Dtsch. Wetterdienst* No. 22, p. 65 (1956).
8. W. KERTZ, *Atmosphärische Gezeiten* in *Encyclopedia of Physics* Vol. 48. Springer, Berlin (1957).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 202–206. Printed in Great Britain.

ATMOSPHERIC WATER VAPOR MEASUREMENTS AT 6–7 MICRONS FROM A SATELLITE

F. MÖLLER

Meteorologisches Institut der Universität München, Amalienstr. 52/III München 13, Germany

(*Received* 20 *October* 1960)

Measurements from an artificial satellite of the long wavelength (terrestrial) radiation emitted and reflected from the earth and its atmosphere should offer some insight on the vertical distribution of temperatures and on the stratification of the radiating gases[1]. The intensity of the total spectrum from about 5 to more than 100μ will not be very useful for his purpose. It is necessary to separate narrow spectral bands by means of filters or even spectrometers.

As a partial problem, we will investigate the kind of information yielded by a measurement within the water vapor spectrum—either in the 6·3μ band or somewhere in the rotation band. We will not enter here into instrumental details but simply assume that we have at our disposal a device to measure the radiation between 6 and 7μ. Then the radiation shall be computed which within this wavelength band is emitted by the atmosphere to a suitable radiation receiver outside the atmosphere.

The atmosphere is considered to be horizontally stratified in parallel and plane layers. The curvature of the earth's surface as well as of the radiation beams is neglected. In this spectral region only the water vapor must be taken into consideration as a radiating gas. The water vapor mass may be defined by

$$u = \int_z^\infty \rho_w \frac{p}{p_0}\, dz \qquad (1)$$

where ρ_w = water vapor density, p = air pressure and $p_0 = 1000$ mb. Herein, the reduction by the factor p/p_0 takes into consideration the effect of pressure broadening of the spectral lines. The radiation is given by

$$I = \int_0^v B_{\Delta\lambda}(T(u)) \frac{dz(u)}{du}\, du + \int_v^\infty B_{\Delta\lambda}(T_s) \frac{dz(u)}{du}\, du \qquad (2)$$

where $B_{\Delta\lambda}$ = Planck's radiation in the range 6–7μ, $T(u)$ = air temperature, $z(u)$ = the emissivity of water vapor in the specified spectral range and T_s = temperature of the earth's surface.

From Daws' data for $z(u)$, a generally valid absorption function and a generalized absorption coefficient as function of the wave number are given.[2] The wavelength band from 6 to 7μ can be approximated by two coefficients, 1·80 and 2·40, each of which is valid in two partial ranges that are again subdivided into two smaller ranges (Fig. 1).[2] With these absorption coefficients, the absorption functions listed in Table 1 for the wavelength range 6–7μ were obtained.

Planck's radiation in the wavelength range 6–7μ is given in Table 2.

For our calculations, only rough basic data were used. Undoubtedly, there may exist better absorption measurements or calculations than Daw's. This absorption function, however, takes into account the general line structure of the absorption band by its shape and the effect of pressure broadening on the line widths by the factor p/p_0 (equation 1). The exact line shape, its variation with p and T, and the change of the line intensity with T are not considered. In addition, the change of Planck's function with T between the wavelength range $\Delta\lambda$ was not considered. This omission is not so serious

Table 1. Absorption function $\varepsilon(u)$ of water vapor between 6 and 7 μ

u (cm H_2O)	$\varepsilon(u)$ (%)	u (cm H_2O)	$\varepsilon(u)$ (%)	u (cm H_2O)	$\varepsilon(u)$ (%)
10^{-7}	1·55	4×10^{-4}	29·7	10^{-2}	78·5
10^{-6}	2·90	6×10^{-4}	34·8	$1·5\times10^{-2}$	83·8
4×10^{-6}	4·80	8×10^{-4}	38·8	2×10^{-2}	87·2
10^{-5}	6·72	10^{-3}	42·0	3×10^{-2}	91·5
2×10^{-5}	8·9	$1·5\times10^{-3}$	48·2	4×10^{-2}	94·1
4×10^{-5}	11·3	2×10^{-3}	52·8	6×10^{-2}	96·1
6×10^{-5}	13·8	3×10^{-3}	59·4	8×10^{-2}	97·4
8×10^{-5}	15·5	4×10^{-3}	64·1	10^{-1}	98·1
10^{-4}	17·0	6×10^{-3}	70·7	$1·5\times10^{-1}$	99·4
2×10^{-4}	22·5	8×10^{-3}	75·1	2×10^{-1}	99·6
				3×10^{-1}	99·8

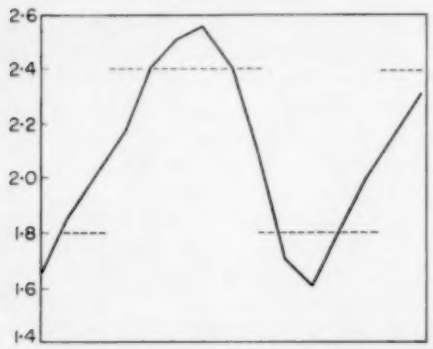


Fig. 1. Generalized absorption coefficient between 6 and 7μ after Daw[2] (solid line) and simplified assumption (dashed line).

because of the distribution of the absorption coefficients over four partial wavelengths. The following calculations give an estimate of the changes of radiation which can be expected when the structure of the radiating atmosphere varies. A more detailed analysis of the spectral structure will not yield essentially different results because the influence of the atmospheric stratification is of a higher order of magnitude.

The integration of equation (2) has been performed graphically in a diagram which has the abscissa u (scaling ε) and the ordinate T (scaling $B_{\Delta\lambda}$). Some examples were calculated, the result given in Table 4.

Example 1

(a) Temperature distribution of the standard atmosphere (troposphere 6·5°C/km, stratosphere isothermal at −55°C; cloudless; 100 per cent rel. humidity, below −20°C related to vapor pressure over ice. Above the tropopause further decrease of the frost point by 7°C/km until at −83°C, which corresponds to about 1 per cent rel. humidity over ice at −50°C (1·8 per cent at −55°C). Above 20 km the same amount of water vapor was assumed as between 15 and 20 km, $\Delta u = 1·55\times10^{-4}$ cm—the total water vapor content of the stratosphere being $4·10^{-4}$ cm.

(b) The same temperature distribution as the standard atmosphere, relative humidity in the troposphere 50 per cent and above the tropopause decrease of the frost point to −83°C.

(c) Same conditions as above, but 10 per cent within the troposphere.

Table 2. Planck's radiation within the wave length range 6–7 μ

T(°K)	$B_{\Delta\lambda}$ (cal/cm² min)	T(°K)	$B_{\Delta\lambda}$ (cal/cm² min)
180	$2·22\times10^{-4}$	250	$6·68\times10^{-3}$
190	$3·65\times10^{-4}$	260	$9·38\times10^{-3}$
200	$7·46\times10^{-4}$	270	$1·281\times10^{-2}$
210	$1·261\times10^{-3}$	280	$1·714\times10^{-2}$
220	$2·124\times10^{-3}$	290	$2·250\times10^{-2}$
230	$3·11\times10^{-3}$	300	$2·898\times10^{-2}$
240	$4·64\times10^{-3}$		

Example 2

(a) Temperature distribution of a tropical atmosphere as in Table 3. Above 20 km further increase of temperature until $^{-}45^{\circ}C$ at $u=0$; cloudless; in the troposphere 100 per cent relative humidity over water, below $-20^{\circ}C$ over ice; decrease of frost point in the stratosphere until $-83^{\circ}C$; above 20 km the same water vapor content as between 17 and 20 km. Total water vapor content of the stratosphere $2{\cdot}8\times10^{-5}$ cm.

(b) Same tropical atmosphere, but relative humidity in the troposphere 50 per cent.

(c) Identical tropical atmosphere as in 2(a), but relative humidity in the troposphere 10 per cent.

Table 3. Assumed temperature distribution of tropical atmosphere

h (km)	T (°C)	h (km)	T (°C)
0	26	12	−44
2	19	14	−62
4	10	16	−80
6	−1	18	−71
8	−13	20	−65
10	−27		

Example 3

(a) Polar atmosphere: at the surface $T=0^{\circ}C$, lapse rate of $^{-}6{\cdot}5^{\circ}C$/km up to 10 km, isothermal stratosphere at $-65^{\circ}C$. Relative humidity of the troposphere 100 per cent, above the tropopause decrease of the frost point by $8^{\circ}C$/km until $^{-}83^{\circ}C$. Above 20 km the same vapor amount as between 15 and 20 km. Total water vapor content of the stratosphere $1{\cdot}4\times10^{-4}$ cm.

(b) The same temperature as in 3(a), relative humidity in troposphere 50 per cent.

(c) Same temperature as in 3(a), relative humidity in troposphere 10 per cent.

The calculated radiation quantities are given in Table 4. Also the radiation temperature is given, which is defined as the temperature of a black radiator which in the special region is in equilibrium with the radiation received from the atmosphere below it.

The three atmospheres given in Table 4 cover a temperature range of 29°C (the temperatures at 8 km being −13°C, −37°C and −52°C, respectively). The radiation temperatures of the same three atmospheres, however, cover only a range of 6·1°C (100 per cent), 6·2°C (50 per cent) and 5·4°C (10 per cent). These small differences are based upon the fact that with the same relative humidity and the same temperature, we have the same water vapor amount. When the vertical stratification of temperature is approximately the same, then the lapse rate for the warm and cold atmosphere is approximately the same above a point with the same temperature T; consequently we have on top of the same temperature the same water vapor mass (when we neglect the $p/p_{\circ}$ effect). Thus, radiation of a given temperature is emitted from the same optical depth in all atmospheres with the same relative humidity. By considering the above, the small variations in the outgoing radiation become understandable. A change of the relative humidity, however, has a much stronger influence. A decrease from 100 to 10 per cent increases considerably the optical depth of the radiation sources and therefore increases the equilibrium temperature T_I in the three atmospheres by 16·3°C, 15·3°C and 17·0°C, respectively. Thus, the radiation within the vibration rotation bands at 6μ can be used as a hygrometer rather than as a thermometer.

Example 4

(a) and (c). The same temperature and humidity distributions were assumed as in examples 1(a) and (c), standard atmosphere with 100 per cent and 10 per cent humidity respectively. Within the very dry stratosphere, however, a layer 500 m thick was assumed to have water vapor saturation over ice between 13·5 km and 14·0 km. The humidity measurements of Barrett *et al.* point to the occurrence of such moist stratospheric layers.[8]

It is inferred from the above that the emissivity of the stratosphere is increased whereas the corresponding radiation temperature T_I is decreased in case 4(a) from −41·6 to −42·0°C and in case 4(c) from −26·3 to −28·4°C.

Table 4. Radiation[1] of the atmosphere between 6–7μ directed vertically upward (10^{-3} cal cm^{-2}min$^{-1}\mu^{-1}$) and radiation temperature T_I

Atmospheric model	(a) 100%		(b) 50%		(c) 10%	
	I	T_I	I	T_I	I	T_I
1. Standard atm.	3·30	–41·6	3·83	–37·9	5·93	–26·3
2. Tropical atm.	3·75	–38·4	4·50	–34·0	6·89	–22·1
3. Polar atm.	2·92	–44·5	3·48	–40·2	5·68	–27·5

Similar calculations can be made for the assumption of clouds which exist within the atmosphere and which may be considered as black body radiators. If, however, the upper boundaries of the clouds in the various climatic regions differ in height but have equal temperatures, then the radiation amount will differ only slightly.[4]

Some more insight into the behaviour of the atmosphere may be expected from measurements of the angular dependence of the radiation. When the radiation strikes the receiver at normal incidence and is inclined by the angle z to the nadir direction, we have—between a special radiating layer in the atmosphere and the receiver—an absorbing water vapor mass which is larger ($m = \sec z$). Hence, the total amount of radiation will be smaller because the temperature in the atmosphere decreases upward. Instead of equation (2) we have:

$$I(z) = \int_0^v B_{\Delta\lambda}(T(u))\,d\varepsilon(mu) + \int_v^\infty B_{\Delta\lambda}(T_a)\,d\varepsilon(mu) \tag{3}$$

Again, the curvature of the earth's atmosphere is neglected. We calculated the radiation $I(z)$ and the radiation temperatures T_I for the three 50 per cent-atmospheres 1(b), 2(b) and 3(b) under the nadir angles $z = 0°$ ($m = 1$), 60° ($m = 2$), 71° ($m = 3$), 79° ($m = 5$). Table 5 gives the results.

The radiation temperatures decrease clearly with increasing angle z. The decrease is nearly uniform in all three cases. It is from $m = 1$ to 5 in the tropical, standard and polar atmospheres 10·2°C, 9·0°C and 10·1°C, respectively. In the standard atmosphere, the decrease is somewhat smaller. This may be a sign that higher stratospheric temperatures exist within this atmosphere.

CONCLUSIONS

The numerical calculations of the atmospheric radiation in the center of the 6·3μ band of water vapor demonstrate that only very small differences occur between warm and cold atmospheres (difference 40°C) when the relative humidity remains unchanged. Changes of humidity between 100 per cent and 10 per cent cause differences to occur within a factor of 3; however, the assumption which was made of an atmosphere with 100 per cent or 10 per cent humidity throughout all heights probably does not occur in nature. The evident reason for these small changes of the radiation is due to the fact that the (saturation) vapor pressure changes simultaneously with the temperature and therefore a given temperature corresponds to the same optical depth with the same relative

Table 5. Terrestrial radiation between 6 and 7μ arriving under nadir angle z and corresponding radiation temperature

Atmospheric model	$m=1$	$m=2$	$m=3$	$m=5$
Tropical atm.	$4·5\times10^{-3}$ –34·0°C	$3·75\times10^{-3}$ –38·4°C	$33·8\times10^{-3}$ –41·0°C	$2·96\times10^{-3}$ –44·2°C
Standard atm.	$3·83\times10^{-3}$ –37·9°C	$3·22\times10^{-3}$ –42·2°C	$2·95\times10^{-3}$ –43·8°C	$2·63\times10^{-3}$ –46·9°C
Polar atm.	$3·48\times10^{-3}$ –40·2°C	$2·88\times10^{-3}$ –44·8°C	$2·61\times10^{-3}$ –47·1°C	$2·27\times10^{-3}$ –50·3°C

humidity. The same conditions, as mentioned above, exist in the other parts of the water vapor spectrum. Hence, they make impossible any conclusion on the temperature distribution in the atmosphere which might be drawn from measurements of the water vapor radiation by a satellite. These restrictions are not valid for the radiation of carbon dioxide which has a vertical distribution independent of the temperature.

REFERENCES

1. J. I. KING, *Scientific Uses of Earth Satellites* p. 133. University of Michigan Press, Ann Arbor (1956).
2. H. A. DAW, *Transmission of Radiation Through Water Vapor*, Tech. Rep. No. 10 Contr. AF 19 (122)-392 Dept of Phys., Univ. of Utah (1956).
3. E. W. BARRETT, R. HERNDON, JR. and H. J. CARTER, *Tellus* **2**, 302 (1950).
4. G. C. SIMPSON, *Mem. Roy. Met. Soc.* **2**, No. 16 (1928).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 207-212. Printed in Great Britain.

OPTIMUM RESOLVING POWER FOR AN ULTRAVIOLET SPACE TELESCOPE

R. J. DAVIS and E. A. GODFREDSEN
Smithsonian Astrophysical Observatory, Cambridge 38, Mass.

(*Received* 21 *October* 1960)

Abstract—The apparent photographic magnitude to which a space telescope (Celescope) can detect stars of various spectral types has been determined by the use of theoretical stellar atmospheres. This information, together with available stellar statistics, was used to obtain an estimate of the density of stars expected to be seen as a function of position on the celestial sphere. On the assumption of statistical homogeneity, it was found that with a sensitivity threshold of the system of 10 photons cm^{-2} sec^{-1}, and a resolving power 1′ of arc, one could expect about 4 per cent of the observed stars to be unresolved doubles. Depending on the wavelength region, from 3×10^5 to 6×10^5 stars are expected to be seen with Celescope.

1. INTRODUCTION

Artificial satellite techniques now make possible extensive astrophysical observations from above the earth's atmosphere. Of most immediate interest to non-solar astrophysicists is a study of the "astrophysical ultraviolet," the spectral region extending from 912 Å to 3000 Å. Theories of the atmospheres of hot stars depend heavily upon measurements made in this region, which cannot be studied from the earth's surface. Interstellar matter can be more effectively studied in this region than in presently accessible ones. Hot atmospheres surrounding cool stars may be discovered and analysed.

The extreme ultraviolet, from below 50 Å to about 912 Å, is probably blocked by interstellar and interplanetary matter, so that only the sun can be effectively studied here. This region, too, is of extreme interest; we must look for windows in this obscuration. When the difficult techniques of manufacturing X-ray telescopes are mastered, the region below 30 Å should yield valuable data.

The Smithsonian Astrophysical Observatory is now designing some of the first space telescopes (Celescopes) to be used for a survey of the sky in the astrophysical ultraviolet. Our immediate aim is to conduct an all-sky survey of stellar and interstellar radiation in the 1000 to 3000 Å range. Our detector is to be a television system, with an angular resolving power of about one minute of arc. Present theories of stellar atmospheres and the few existing preliminary rocket observations have led us to believe that in this wavelength range we will be detecting primarily hot stars, of spectral type *A*0 and earlier. The ultraviolet nebulosities, so conspicuous in the rocket observations, will become less evident because of our higher angular resolving power.

In order to improve our theories of stellar atmospheres, we desire a statistically significant sample of hot main-sequence stars. By setting as our goal a sensitivity of 10 photons sec^{-1} cm^{-2}, we expect to be able to measure the fluxes, in several wavelength intervals, of more than 10^5 stars, obtaining a significant sample of stars of types 0, *B* and *A*. With our contemplated optical system (an 8-in. or slightly larger reflecting telescope), and integrating time of 10 secs, the detector should have a quantum efficiency of 10^{-5}.

In this paper, we discuss the effects of resolving power on our observational results. We are particularly interested in the number of unresolved stars we can expect to observe, since the value of our observational results will depen strongly upon identification of observed ultraviolet sources with known objects.

We do not consider here the effects of interstellar extinction on the number of stars we

expect to observe. According to Code[14], the inverse wavelength dependence of extinction caused by interstellar dust is beginning to level off near the lower end of the observed range, at about 3200 Å. He predicts that the dust will cause even less absorption at 1000 Å than at 3000 Å. Since we have no better theory to offer, we have not taken absorption by interstellar dust into account in our computations.

In addition to the dust, however, the interstellar gas can be expected to cause continuous absorption. The strongest contributors to this absorption are expected to be sodium, below 2413 Å; aluminium, below 2072 Å; calcium, below 2029 Å; nickel, below 1624 Å; magnesium, below 1622 Å; iron, below 1570 Å; silicon, below 1521 Å; sulphur, below 1197 Å; phosphorus, below 1175 Å; carbon, below 1101 Å and ionized calcium, below 1044 Å. We cannot determine the amount of this absorption because of poor cross section and abundance data, and therefore we have not been able to include it in our computations.

2. USE OF MODEL ATMOSPHERES

Since the Celescope is intended to observe the radiation of stars in the far ultraviolet part of the spectrum, most of the information it provides will concern the hotter stars. For these stars, the maximum of the intensity curve falls within the frequency limits being investigated. The contribution of the ultraviolet part of the spectrum of stars of spectral type *G* and later is so small that relatively few nearby stars of this kind are expected to be observed.

This factor is important in the attempt to deduce the expected intensity of spectra of early-type stars. Because we are concerned primarily with stars with masses greater than one solar mass, only the conditions of radiative and mechanical equilibrium must be fulfilled in the atmospheres of the star in order to approximate the physical conditions of a real star. By applying the physical theory of the interaction between radiation and matter, one may compute a theoretical spectrum. When this computed spectrum matches an observed stellar spectrum we may conclude that the model represents a true stellar atmosphere. Then we can use this model to extrapolate into the as yet unobserved part of the spectrum. The validity of this extrapolation will be determined by a comparison with the observations obtained from telescopes orbiting above the earth's atmosphere. Table 1 lists twelve model atmospheres that we have used to make such an extrapolation.

The computed spectral distribution differs considerably from that of a black body at the effective temperatures of the respective stars, primarily because of the absorption edges of various elements in the stellar atmospheres. A typical curve is shown in Fig. 1. At a wavelength of 912 Å, the limit of the Lyman series, the absorption of the star's radiation by the hydrogen in the stellar atmosphere causes a

Table 1. Model atmospheres used to obtain apparent photographic magnitudes

Spectral class	T_e (°K)	log. g	Year published	Author
05 V	44,600	4·2	1955	Underhill[1]
06·5	30,700*	4·2	1959	Saito and Uesugi[2]
09 V	37,450	4·45	1956	Traving[3]
*B*0	32,800	4·8	1955	Traving[4]
*B*2·5	16,800*	3·8	1957	Underhill[5]
*A*0	9,200	4·3	1955	Hunger[6]
*A*3	8,900	4·5	1956	Osawa[7]
*A*9	7,560	4·5	1956	Osawa[7]
*F*2·5	6,720	2·8	1950-51	Vitense[8]
*G*0	5,930	2·8	1950-51	Vitense[8]
*G*1·5 (sun)	5,785	4·45	1959	Aboud *et al.*[9]
				Rense[10]
*G*8·5	5,305	2·8	1950-51	Vitense[8]

*Boundary temperature.

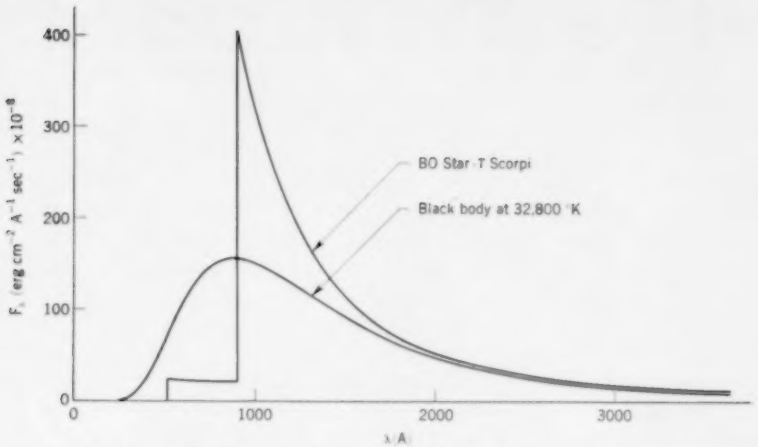


Fig. 1. Theoretical spectral energy curve of a B0 star compared with that of a black body at 32,800°K.

large discontinuity in the spectral energy curve. The relative importance of various absorption edges depends on the spectral class of the star. Thus, for the sun, rocket observations show a large decrease in the continuum below a wavelength of about 2000 Å. This decrease is caused by absorption by aluminium and calcium atoms in the sun's atmosphere beyond their series limits. This rather recent discovery suggests that new observations in the ultraviolet part of the spectrum of other stars may show features that have not been predicted by the model atmospheres so far constructed. However, until such observations are available, the theoretically computed model atmospheres can serve as a guide in the attempt to estimate the amount of energy coming from early-type stars within the frequency ranges to be detected by Celescope.

3. STATISTICAL ANALYSIS

In a remote observing instrument such as Celescope, the photographic plate is replaced by a television scanning system. The ability of the system to resolve individual stars will, therefore, depend on the electronic circuitory of the instrument and will differ markedly from that of a photographic plate. In order to estimate the percentage of the unresolved stars, one must first be able to estimate the density of the star field as a function of position in the sky. This has been done in the following manner.

The distribution of stars on the celestial sphere is far from uniform, and a strong concentration of stars occurs toward the galactic equator. Therefore, it is desirable to know the density of stars as a function of galactic co-ordinates. This information appears in a table published by Allen (Ref. 11, p. 213) which lists the number of stars per square degree brighter than a given apparent photographic magnitude as a function of galactic latitude.

Since the temperature of a star is so important in determining the amount of energy that can be detected by Celescope, we also want to know the distribution of stars as a function of stellar type. Although this distribution is readily available for stars as a whole, it has not as yet been listed as a function of galactic latitude. It is quite relevant since the percentage of hot young stars decreases as the galactic latitude increases. In order to obtain this distribution by galactic latitude, we classified all the stars in the Boss General Catalogue[12] according to their spectral type and position. This task was greatly facilitated by the fact that this catalogue has been put on magnetic tape for use in a high speed computer (IBM-704). In addition, we used a list of 1270 O-*B* stars.[13] If a given star

appeared in both sources, we used the spectral type as given by Morgan *et al.*[13].

By combining the density of stars and their distribution by spectral type in a given magnitude interval, we can obtain the number of stars per square degree brighter than a given magnitude as a function of spectral class and position on the celestial sphere. We can then calculate the number of stars per square degree to be expected on the basis of model atmospheres described above. From the spectral energy curve, we can construct a curve giving $q(\lambda)$, the number of photons $sec^{-1}cm^{-2}A^{-1}$, as a function of wavelength. The area beneath the $q(\lambda)$ curve in the frequency range under consideration must be equal to or exceed Q, the number of photons $sec^{-1}cm^{-2}$ that can be detected by the telescope and detector system. This equation,

$$\int_{\lambda_1}^{\lambda_2} q\,(\lambda)\,d\,\lambda \geqslant Q \qquad . \; . \; . \quad (1)$$

acts to determine the vertical scale of the $q(\lambda)$ curve and consequently the scale of the spectral energy curve. Once this scale has been determined, the total area under the spectral energy curve gives the total flux, F, coming from the star. By definition, a star whose flux is $2{\cdot}27 \times 10^{-5}$ ergs cm^{-2} sec^{-1} above the earth's atmosphere has an apparent bolometric magnitude of zero. The limiting apparent bolometric magnitude to which the telescope can detect stars of various special types is then given by the equation:

$$m_{bol} = 2{\cdot}5 \log \left(\frac{2{\cdot}27 \times 10^{-5} \text{ ergs cm}^{-2} \text{ sec}^{-1}}{F} \right) \qquad . \; . \; . \quad (2)$$

By applying the bolometric correction and color index given by Allen (Ref. 11, p. 183), we can obtain the limiting apparent photographic magnitude to which we can expect to detect stars of a given spectral type.

A combination of all the information just described can, in principle, be used to calculate the expected number of stars per square degree as a function of position in the sky. This calculation, which was made feasible by the use of the IBM-704, allowed us to estimate the percentage of unresolved stars by assuming a statistically random distribution of stars throughout the 2° square field of view of the Celescope. The probability, $dp(r)$, that a star's closest neighbor will be between the limits of r and $r+dr$ is then given by the equation

$$dp(r) = (2r/r^2_{avg}) \exp\left[-(r/r_{avg})^2\right] dr \qquad . \; . \quad (3)$$

Thus the probability that a star's closest neighbor lies inside a circle of radius r is given by the equation:

$$P(r) = \int_0^r p(r^1)\,dr^1 = -\int_0^r \exp\left[-(r^1/r_{avg})^2\right] dr^1$$

$$= 1 - \exp\left[-(r/r_{avg})^2\right] \qquad . \; . \; . \quad (4)$$

The quantity r_{avg} is a length defined by the relation, $n_{avg} = 1/(\pi\, r^2_{avg})$, where n_{avg} is the expected number of stars per unit area. The total number of unresolved stars is then given by the product of $P(r)$ and N, the total number of stars considered. Half this number gives the number of unresolved pairs of stars. This method of analysis yields estimates of the number of stars per square degree and of the percentage of unresolved stars, both as functions of position on the celestial sphere.

4. APPLICATION TO CELESCOPE

We have applied the method of analysis described above, using the conditions to be expected for Celescope, and have obtained results for five different wavelength ranges: 1050–1230 Å, 1230–1400 Å, 1400–1600 Å, 1600–2200 Å and 2200–3000 Å. Fig. 2 shows the curves for the five different spectral ranges of Celescope's limiting apparent photographic magnitude as a function of stellar type. These curves were calculated on the assumption of a minimum sensitivity threshold of 10 photons sec^{-1} cm^{-2}. The shape of the curves is explained by the fact that later spectral types usually correspond to stars of lower tempera-

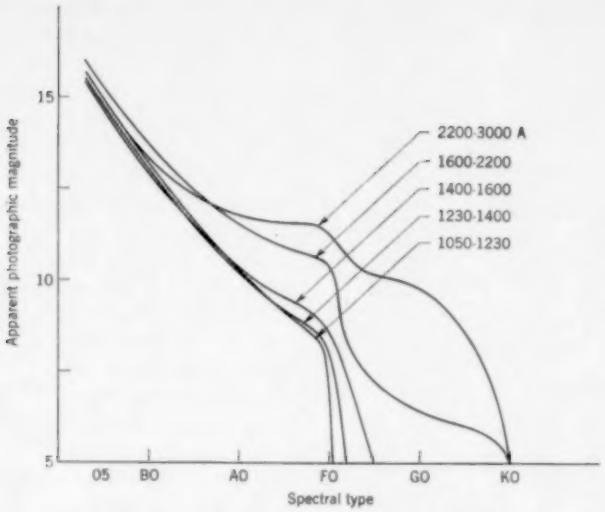


Fig. 2. Apparent photographic magnitude of stars detectable by Celescope, as a function of spectral type.

tures and, consequently, a shift of the maximum of the spectral energy curve away from the frequency ranges being used.

The results from these curves were combined with Allen's[11] table of stars per square degree. We divided the celestial sphere into 288 sections: twenty-four divisions of right ascension and twelve divisions of declination. We determined the average number of stars per square degree expected in each section. The total number of stars expected to be observed in each of the five wavelengths ranges is shown in Table 2. The number of unresolved stars expected is also given for resolving powers of 10, 30 and 60″ of arc. A television scanning system of 500 lines per 2° field will give a resolution of about 60″ of arc. The actual number of unresolved stars may somewhat exceed the figures given because in some sections of the sky the distribution of stars is not statistically random. In the galactic plane there is a tendency for stars to be found in groups or associations.

Figure 3 gives an idea of the actual number of stars per square degree to be expected at various galactic latitudes, and Fig. 4 gives a more detailed description. In some regions the average number of stars per square degree exceeds 200; certain fields have as many as 1000 stars per field. On the other hand, areas within 50° of the galactic poles usually have less five stars per square degree and often only one or

Table 2. Numbers of observable and of unresolved stars

Wavelength range (Å)	Unresolved stars						Total number of observable stars
	($r_{min}=10''$)		($r_{min}=30''$)		($r_{min}=60''$)		
	number	%	number	%	number	%	
1050–1230	700	0·1	6,200	1·0	24,000	3·9	610,000
1230–1400	1,240	0·2	10,000	1·5	40,000	5·9	680,000
1400–1600	250	0·1	2,200	0·8	8,600	3·0	290,000
1600–2200	250	0·1	2,200	0·8	8,700	3·1	280,000
2200–3000	420	0·1	3,700	1·1	15,000	4·3	350,000

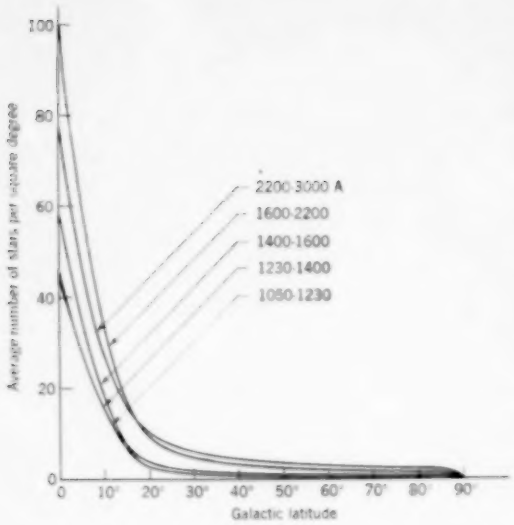


Fig. 3. Average number of stars per square degree, as a function of galactic latitude.

two. This wide range of star densities must be considered in our attempts to develop an efficient electronic system for delivering information from the telescope to the observer on the earth.

5. CONCLUSION

Table 2 indicates, as we would expect, that the number of unresolved double stars is proportional to the square of the resolving power of the telescope. If we set our resolving power at 1′ of arc as planned, we can expect to identify correctly about 96 per cent of our observed stars. To identify 99 per cent of our stars correctly, we must double the resolving power. With the fixed linear resolving power and field of view of the Uvicon television system, this would mean doubling the focal length of the system, to 48 in., and quadrupling the time required to scan the entire sky. On the other hand, if we are willing to accept 90 per cent correct identification, we can shorten the observing time to somewhat under 1 month, compared to 4 months for 96 per cent correct identification, but must reduce the focal length to 8 in., entailing rather difficult optical systems.

As Table 2 also indicates, the percentage of unresolved double stars is not very sensitive to the total number of stars; indeed, the differences in distributions of the various spectral types obscure the small effect that does exist. We conclude that, unless a better television system can be produced, 1′ of arc is the optimum resolving power. 30″ resolving power would be preferable if we could keep the same field of view, 2 degrees square. We also conclude that our proposed threshold of 10 photons cm^{-2} sec^{-1} is adequate to provide a statistically significant sample of stars if our theories are correct to within a factor of 10. With the proposed Uvicon television system, we can perhaps increase this sensitivity by an additional order of magnitude by increasing the integrating time.

REFERENCES

1. A. B. Underhill, *Domin. Astrophys. Observat.* **8,** 357 (1955).
2. S. Saito and A. Uesugi, *Contr. Inst. Astrophys. and Kwasan Obs.* No. 78 (1959).
3. G. Traving, *Z. Astrophys.* **41,** 215 (1956).
4. G. Traving, *Z. Astrophys.* **36,** 1 (1955).
5. A. B. Underhill, *Domin. Astrophys. Observat.* **10,** 357 (1957).
6. K. Hunger, *Z. Astrophys.* **36,** 42 (1955).
7. K. Osawa, *Astrophys. J.* **123,** 513 (1956).
8. E. Vitense, *Z. Astrophys.* **28,** 81 (1950–51).
9. A. Aboud, W. E. Behring and W. A. Rense, *Astrophys. J.* **130,** 381 (1959)
10. W. A. Rense, *Sky and Telescope* **17,** 502 (1958).
11. C. W. Allen, *Astrophysical Quantities,* Athlone Press, London (1955).
12. B. Boss, *General Catalogue of 33,342 Stars for the Epoch 1950.* Carnegie Institute of Washington (1937).
13. W. W. Morgan, A. D. Code and A. E. Whitford, *Astrophys. J. Suppl. Ser.* **14,** 41 (1955).
14. A. D. Code, *Astr. J.* **65,** 278 (1960).

(Facing page 212)

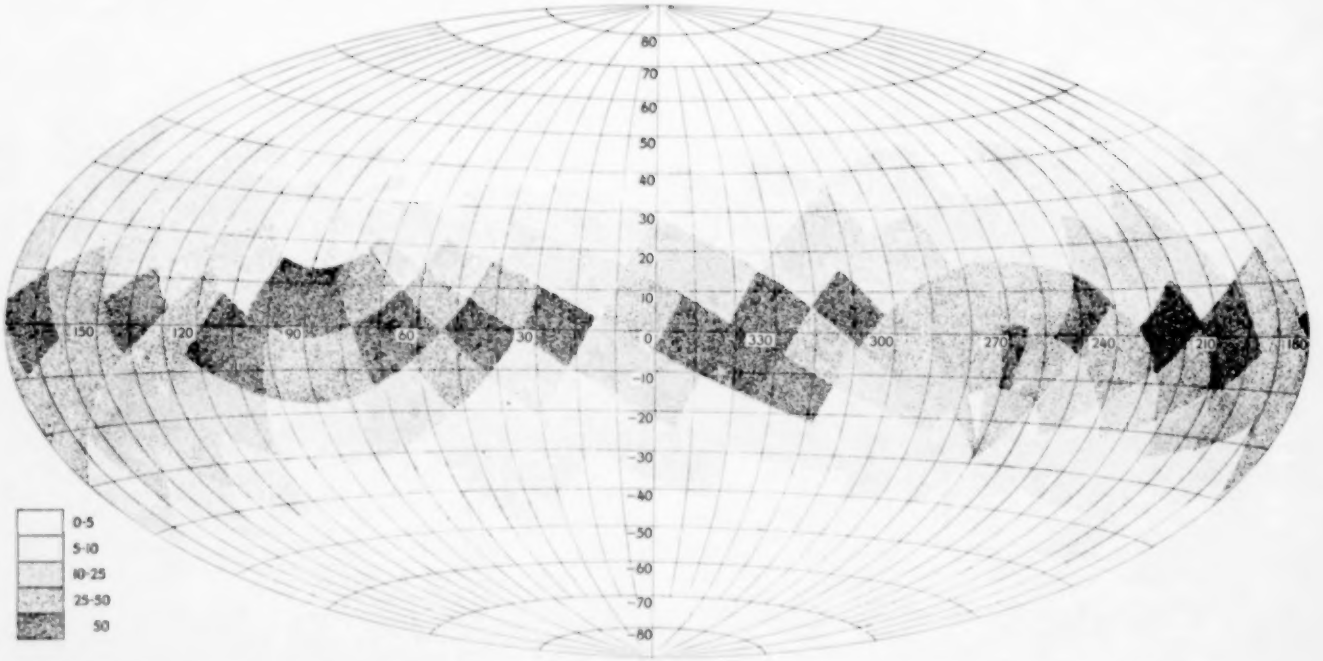


Fig. 4. Expected distribution of average number of stars per square degree on the celestial sphere in wavelength region of 2200–3000 Å.

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 213–219. Printed in Great Britain.

LOCAL ELECTRON DENSITIES DEDUCED FROM THE FARADAY FADING OF SATELLITE TRANSMISSIONS USING MEASUREMENTS DURING TWO CONSECUTIVE TRANSITS

L. LISZKA
Kiruna Geophysical Observatory, Sweden

(*Received* 24 *October* 1960)

Abstract—A simplified method for analysis of the Faraday fading measurements is described. With a satellite of circular orbit, both the total electron content between observer and satellite and the direction of the magnetic field in the ionosphere can be calculated. For all other orbits the local electron density must be taken into consideration. From measurements of the Faraday fading frequency during at least two consecutive transits the total electron content and the local electron density can be calculated as a function of satellite altitude on the following assumptions:

(*a*) During the period of time comprising two consecutive transits the relative gradients of local electron density are constant at the same altitudes.

(*b*) The local electron density at altitudes above the *F2* maximum during a period of time comprising two consecutive transits changes as does the electron density within the maximum of the *F2* layer.

The method has been applied to the daytime observations of 1958 δ2 on the 12 and 29 of September 1959.

1. INTRODUCTION

The polarization plane of satellite signals during passage through the ionosphere rotates through an angle which depends on the frequency of signals, the magnetic field and the local electron density along the ray path between satellite and observer. The signal strength observed with a linearly polarized antenna thus changes periodically. From measurements of this fading, the total electron content of the ionosphere and the local electron density at the satellite altitude can be deduced if the ionosphere parameters are known. In order to ascertain these parameters, simultaneous measurements from several stations should be used. Since this involves practical difficulties, other methods requiring measurements from only one station have been evolved.

The simultaneous observations of fading on two different frequencies[1] should be mentioned here.

Faraday rotation can be observed in various ways, i.e. by measurement of the total angle of polarization rotation at a propitious time, and by measurement of the rate of polarization rotation. A full discussion of these two methods and of corrections required for the simplified theory has been presented by Garriott.[2]

In this paper, measurements of the rate of polarization rotation are analysed by means of simplified equations. The analysis was based on a small number of 1958 δ2 transmission records suited for application of the simplified theory. From the measurements of the rate of polarization rotation during at least two consecutive transits, the total electron content and the local electron density were calculated on the basis of certain general assumptions concerning the ionospheric structure above the *F2* maximum.

2. THEORY OF THE ANALYSIS

In passing through the ionosphere, the polarization plane of the satellite signals rotates through an angle.[3]

$$\theta = 2{\cdot}97 \times 10^{-2} \frac{H}{f^2} \cos\phi \int N dr \quad . \quad (1)$$

where

H = magnetic field in the ionosphere in m.k.s. units;

N = electron density per cubic meter;
ϕ = angle between the direction of propagation and the magnetic field in the ionosphere;
f = transmitted frequency.

This formula is based on the following assumptions:

(1) The transmitted frequency is considerably greater than the plasma and electron gyromagnetic frequencies;
(2) The product $H\cos\phi$ does not depend on height in the ionosphere;
(3) The refraction is negligible (for a frequency of 20 Mc/s, this is valid only for small zenith distances);
(4) The formula cannot be used for ϕ between 80 and 100°.

The signal strength observed with a linearly polarized antenna thus changes periodically with the frequency:

$$\nu = \frac{1}{\pi}\frac{d\theta}{dt} \quad . \; . \; . \quad (2)$$

Formulae (1) and (2) will be converted into equations serviceable for analysis of the observed frequency of fading, orbital data for the satellite and given ionospheric data from standard ionograms being given.

The calculations were based on the geocentric co-ordinate system, with the X axis in the plane of the meridian of Greenwich, the Y axis in the meridan 90° East of Greenwich, and the Z axis coinciding with the earth's axis.

Notations used:

$\mathbf{r}_A$ = radius vector of the observer with co-ordinates X_A, Y_A, Z_A;
$\mathbf{r}_s$ radius vector of the satellite with co-ordinates X_S, Y_S, Z_S;
$\mathbf{k}$ = unit vector of the geomagnetic field;
Z = time of the equatorial crossing in UT Greenwich;
W = longitude of the subsatellite point of the equatorial crossing, measured west of Greenwich from 0° to 360°;
τ = time in minutes, calculated from the moment of the last equatorial crossing;
ϕ_s = geographic latitude of the subsatellite point;
λ_s = geographic longitude of the subsatellite point, calculated in the same direction as W;
h = altitude of the satellite above the geoid at the subsatellite point;
ζ = zenith distance of satellite, as seen by the observer;
$\boldsymbol{\rho} = \mathbf{r}_A - \mathbf{r}_s$

It is evident that

$$\tau = t(\mathrm{UT}) - Z \quad . \; . \; . \quad (3)$$

$$\lambda_s = W + \delta\lambda_s(\tau) \quad . \; . \; . \quad (4)$$

The data: $\phi_s(\tau)$ $\delta\lambda_s(\tau)$, $h(\tau)$ Z and W were taken from *Space Track Bull.* and were used for calculations of $\cos\phi \int N dr$ in expression (1). To eliminate the errors caused by the time Z values given in *Space Track Bull.* the following correction was used. For each analysed transit the time of closest approach was calculated from the Doppler curve. For the time of closest approach $\tau_{\mathrm{C.A.}}$, it is evident that

$$\boldsymbol{\rho}(\tau) \cdot \dot{\boldsymbol{\rho}}(\tau) = 0 \quad . \; . \; . \quad (5)$$

The numerical solution of equation (5) gives the time $\tau_{\mathrm{C.A.}}$, which can be compared with the observed $t_{\mathrm{C.A.}}$ (UT). The true time of equatorial crossing, Z_{corr} may then be given as

$$Z_{\mathrm{corr}} = t_{\mathrm{C.A.}}(\mathrm{UT}) - \tau_{\mathrm{C.A.}} \quad . \; . \; . \quad (6)$$

This procedure is necessary in order to obtain a correct starting time for calculation of τ.

On the aforementioned assumptions

$$\cos\phi = \frac{\mathbf{k} \cdot \boldsymbol{\rho}}{\rho} \quad . \; . \; . \quad (7)$$

and

$$\int N dr = N_\circ \frac{\rho}{h}(1 - L) \; . \; . \; . \quad (8)$$

where $N_\circ = \int_0^{h_s} N(h)dh$ is the total electron content between satellite and ground over a column of 1 m² base, and $L = \frac{1}{2}\left[\rho^2/h\,(1+h)^2\right]\sin\zeta$.

Formula (1) may be written

$$\theta = 2{\cdot}97 \times 10^{-2} \frac{H}{f^2} N_\circ \frac{\mathbf{k} \cdot \boldsymbol{\rho}}{h}(1 - L) \; . \quad (9)$$

The Faraday fading frequency is then given by

$$\nu(\tau)=\frac{2{\cdot}97\times10^{-2}}{60f^2\pi}H\frac{d}{d\tau}\left[\frac{\mathbf{k}\cdot\boldsymbol{\rho}}{h}(1-L)N_0\right]. \quad (10)$$

where ν is measured in $\sec^{-1}$ and τ in min. For a transmitted frequency $f=20$ Mc/s

$$\frac{1}{c}=\frac{2{\cdot}97\times10^{-2}}{60f^2\pi}=2{\cdot}537\times10^{18} \quad . \quad (11)$$

Expression (10) can be written

$$\nu(\tau)=\frac{1}{c}H\left\{N_0\frac{d}{d\tau}\left[\frac{\mathbf{k}\cdot\boldsymbol{\rho}}{h}(1-L)\right]+\frac{\mathbf{k}\cdot\boldsymbol{\rho}}{h}(1-L)\frac{d}{d\tau}N_0\right\} \quad . \quad . \quad (12)$$

$$=\frac{1}{c}HN_0\frac{d}{d\tau}\left[\frac{\mathbf{k}\cdot\boldsymbol{\rho}}{h}(1-L)\right]+\frac{1}{c}H\frac{\mathbf{k}\cdot\boldsymbol{\rho}}{h}(1-L)\frac{dh_s}{d\tau}N \quad . \quad (13)$$

assuming that the horizontal gradient of N_0 may be ignored. At the latitude of Kiruna this assumption seems valid for all cases in which Faraday fading has been observed. This is indicated by the fact that irregularities in the fading frequency have been found to recur at the same altitude in consecutive transits and by the fact that no significant time variations in f_0F2 were recorded at Kiruna for the days in question.

Using notation

$$\boldsymbol{\Psi}(\tau)=\frac{1}{c}\frac{\boldsymbol{\rho}}{h}(1-L) \quad . \quad . \quad (14)$$

$$\nu(\tau)=\mathbf{H}\cdot\frac{d}{d\tau}\boldsymbol{\Psi}N_0+\mathbf{H}\cdot\boldsymbol{\Psi}\frac{dh_s}{d\tau}N(h_s), \quad . \quad . \quad (15)$$

it is possible to calculate that

$$\frac{d}{d\tau}\boldsymbol{\Psi}=\mathbf{II}-L\mathbf{II}-\frac{\boldsymbol{\rho}}{ch}\frac{dL}{d\tau}=\mathbf{II}^0 \quad . \quad . \quad . \quad (16)$$

where

$$\mathbf{II}=\frac{1}{ch(\tau)}\left[\frac{d}{d\tau}\boldsymbol{\rho}(\tau)-\frac{\boldsymbol{\rho}(\tau)}{h(\tau)}\frac{dh(\tau)}{d\tau}\right]. \quad . \quad (17)$$

The components of vector $\mathbf{II}$ in the defined co-ordinate system can be calculated by using the following formulae for the components of vectors $\boldsymbol{\rho}$ and $\dot{\boldsymbol{\rho}}=d\boldsymbol{\rho}/d\tau$

$$x=X_A-(1+h)\cos\phi_s\cos\lambda_s$$
$$y=Y_A+(1+h)\cos\phi_s\sin\lambda_s$$
$$z=Z_A-(1+h)\sin\phi_s \quad . \quad . \quad . \quad (18)$$

The components of ρ are measured in units of the earth's radius.

$$\dot{\rho}_x=-\frac{dh}{d\tau}\cos\phi_s\cos\lambda_s+(1+h)\left[\frac{d\phi_s}{d\tau}\sin\phi_s\cos\lambda_s+\frac{d\lambda_s}{d\tau}\cos\phi_s\sin\lambda_s\right]$$

$$\dot{\rho}_y=\frac{dh}{d\tau}\cos\phi_s\sin\lambda_s-(1+h)\left[\frac{d\phi_s}{d\tau}\sin\phi_s\sin\lambda_s-\frac{d\lambda_s}{d\tau}\cos\phi_s\cos\lambda_s\right]$$

$$\dot{\rho}_z=-\frac{dh}{d\tau}\sin\phi_s-(1+h)\frac{d\phi_s}{d\tau}\cos\phi_s \quad . \quad . \quad (19)$$

For a circular orbit ($dh_s/d\tau=0$), assuming that the fading frequency depends only on the spatial location of the satellite—which means that the geomagnetic field and likewise the total electron content are homogeneous in the relavant region of the ionosphere, and constant during the time of the measurement—it is possible to calculate H_x/H_z, H_y/H_z, H_x/H_y and HN_0 from the equation

$$\Pi^0{}_xH_xN_0+\Pi^0{}_yH_yN_0+\Pi^0{}_zH_zN_0=\nu(\tau) \quad . \quad . \quad . \quad (20)$$

The components $\Pi^0{}_i$ can be calculated from orbital data for those times when $\nu(\tau)$ was measured, and the products H_iN_0 can then be computed by the method of least squares.

In general, the whole of equation (15) must be taken into consideration because both $dh_s/d\tau$ and the local electron density have sufficiently high values.

Using notation

$$\boldsymbol{\Psi}^0=\frac{dh_s}{d\tau}\boldsymbol{\Psi} \quad . \quad . \quad . \quad (21)$$

equation (15) can be written

$$\mathbf{II}^0.\mathbf{H}\,N_0+\boldsymbol{\Psi}^0.\mathbf{H}\,N(h_s)=\nu(\tau) \quad . \quad (22)$$

Assuming the magnetic field to be known—for example, calculated from the eccentric dipole

model—it would be possible to compute both the mean value of the local electron density during the entire transit and the total electron content.

From measurements of the Faraday fading during at least two consecutive transits, the local electron density as a function of a satellite altitude and total electron content can be calculated on the basis of the following assumptions:

(a) During a period of time comprising two consecutive transits the relative gradients of local electron density are constant at the same altitudes.

(b) The local electron density at altitudes above the *F2* maximum during a period of time comprising two consecutive transits changes as does the electron density within the *F2* maximum.

Using notations

$$\mathbf{\Pi}^{\circ}.\mathbf{H}=\alpha \quad . \; . \; . \quad (23)$$

and

$$\Psi^{\circ}.\mathbf{H}=\beta \quad . \; . \; . \quad (24)$$

equation (22) can be written:

$$\alpha N_{\circ}+\beta N(h_s)=\nu \quad . \; . \; . \quad (25)$$

For two consecutive transits with numbers l and k the functions $\alpha(\tau)$, $\beta(\tau)$ and $\nu(\tau)$ must be known in overlapping height intervals. The local electron density is given by

$$N=\frac{\nu}{\beta}-\frac{\alpha}{\beta}N_{\circ} \quad . \; . \; . \quad (26)$$

For the same altitudes in two consecutive transits, owing to assumption (a):

$$\frac{1}{N^{(k)}}\frac{dN^{(k)}}{dh}=\frac{1}{N^{(l)}}\frac{dN^{(l)}}{dh} \quad . \; . \; . \quad (27)$$

and in the approximation:

$$\Delta\left(\frac{\nu}{\beta}\right)^{(k)}-\Delta\left(\frac{\alpha}{\beta}\right)^{(k)}N_{\circ}{}^{(k)}=\frac{N^{(k)}}{N^{(l)}}\Delta\left[\left(\frac{\nu}{\beta}\right)_{(l)}-\left(\frac{\alpha}{\beta}\right)^{(l)}N_{\circ}{}^{(l)}\right] \quad . \; . \quad (28)$$

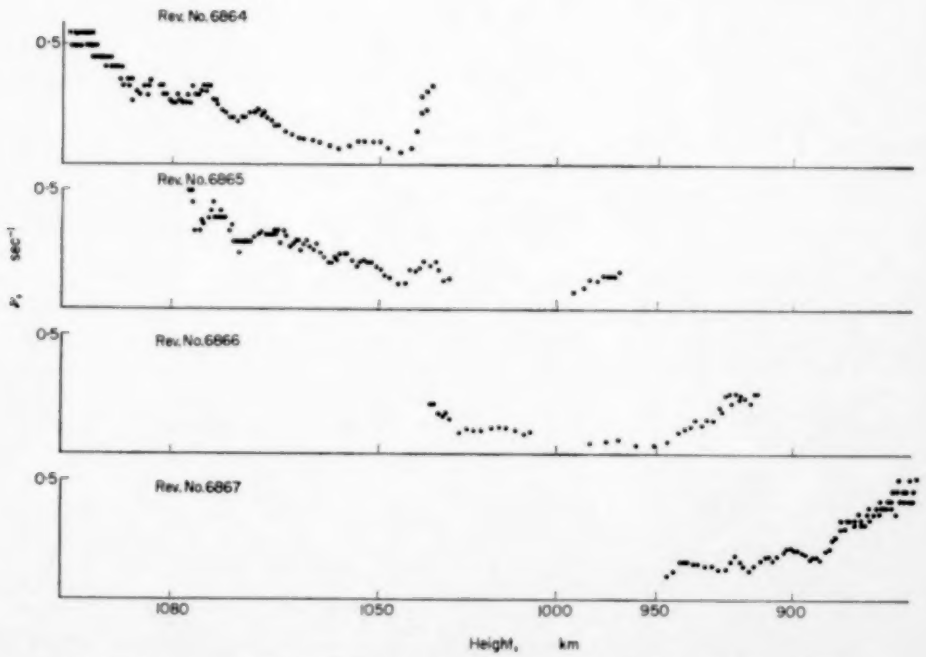


Fig. 1. Frequency of Faraday fading as a function of satellite altitude, 12th Sept. 1959.

Owing to assumption (b):

$$\frac{N^{(k)}}{N^{(i)}} = \frac{N^{(k)}{}_{F2\max}}{N^{(i)}{}_{F2\max}} = \gamma \quad . \quad . \quad . \quad (29)$$

The total electron contents for both transits can be calculated by the method of least squares from the equation

$$\gamma\Delta\left(\frac{\alpha}{\beta}\right)^{(i)} N_o^{(i)} - \Delta\left(\frac{\alpha}{\beta}\right)^{(k)} N_o^{(k)} = \gamma\Delta\left(\frac{\nu}{\beta}\right)^{(i)} - \Delta\left(\frac{\nu}{\beta}\right)^{(k)} \quad . \quad . \quad (30)$$

By applying the calculated values of N_o to expression (26), it is possible to compute the electron-density profiles. The approximations used are adequate for measurement of small zenith distances limited by the actual plasma frequency.

3. RESULTS OF THE ANALYSIS

The discussed method was tentatively employed for measurement of Faraday fading during transits 6864, 6865, 6866, 6867, 7114, 7115 and 7116 on the 12 and 29 of September 1959. The low plasma frequencies (X not much higher than 0·1) allowed use of the measurements for zenith distances up to 55° without correction for dispersion and path-splitting (expression (26) in Ref. 2). The value for the magnetic field was taken from Vestine's tables[4] for 300 km height and Kiruna's co-ordinates, and was extrapolated up to the year 1959. The electron densities at the *F2* maximum and the electron content up to *F2* maximum were computed from $h' - f$ records of the Kiruna ionosounder.

The components of vectors $\mathbf{\Pi}^\circ$ and $\mathbf{\Psi}^\circ$ were calculated for every 15 sec, using formulae (14), (16), (17), (18), (19) and (21) with the aid of the Swedish digital computer, BESK.

With these measurements the Faraday fading frequency can be obtained as a function of time. It can be readily demonstrated that the frequency of fading in the obtained measurements depends not only on the first term of equation (22), but that discrepancies depend on the second term, which is a function of satellite altitude. When the fading frequency is plotted as a function of satellite altitude in several consecutive transits (Fig. 1), certain irregularities are observed at given satellite altitudes. The satellite seems to pass through regions where the electron density is lower or higher than normal rather stable in time. These regions must be greatly extended in the horizontal plane since, during two consecutive transits, the satellite passes through different sectors of the sky.

The two days analysed, the 12th and the 29th September 1959, were free of magnetic disturbances. During the night of 28–29th September moderate magnetic disturbances occurred. The total electron content calculated by means of equation (30) as well as the electron content up to the *F2* maximum are plotted in Figs. 2 and 3 as a function of local time. The determined values for total electron content are, unfortunately, associated with fairly high errors.

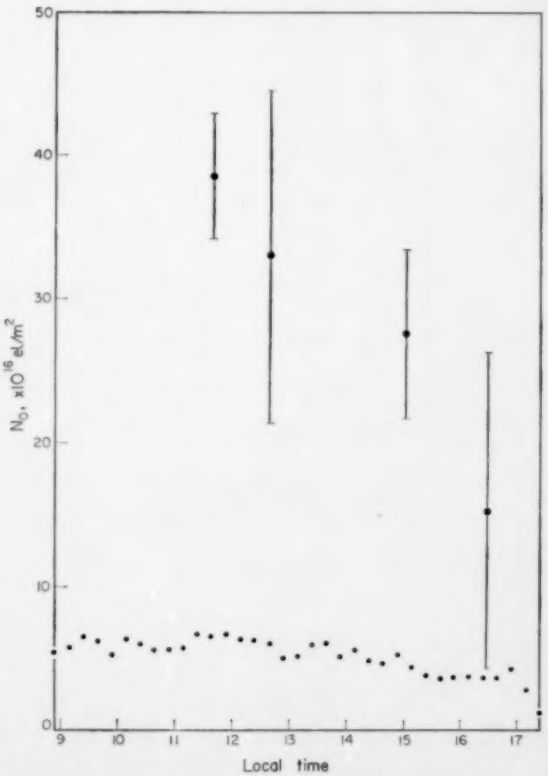


Fig. 2. Electron content up to F2 maximum and calculated total electron content up to satellite altitude, 12th Sept. 1959.

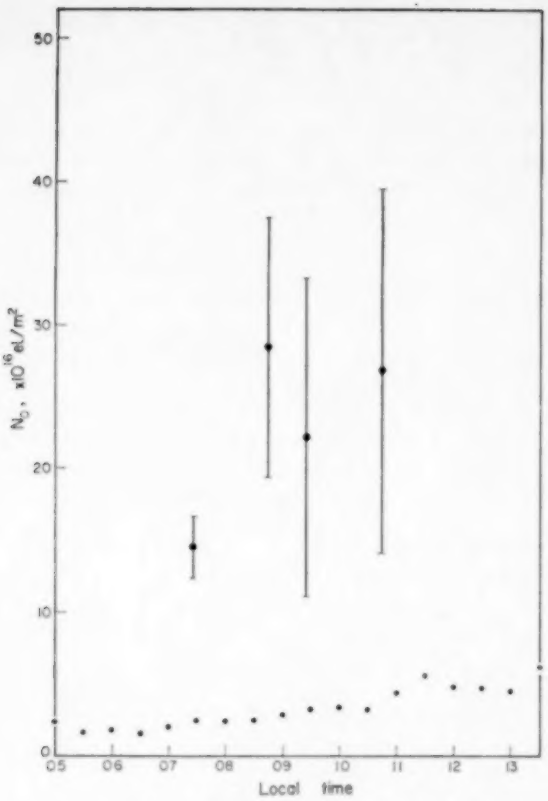


Fig. 3. Electron content up to *F2* maximum and calculated total electron content up to satellite altitude, 29th Sept. 1959.

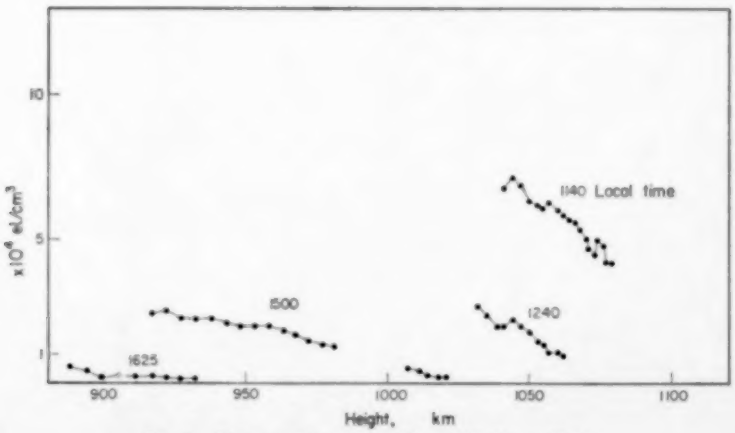


Fig. 4. Electron density profiles, 12th Sept. 1959.

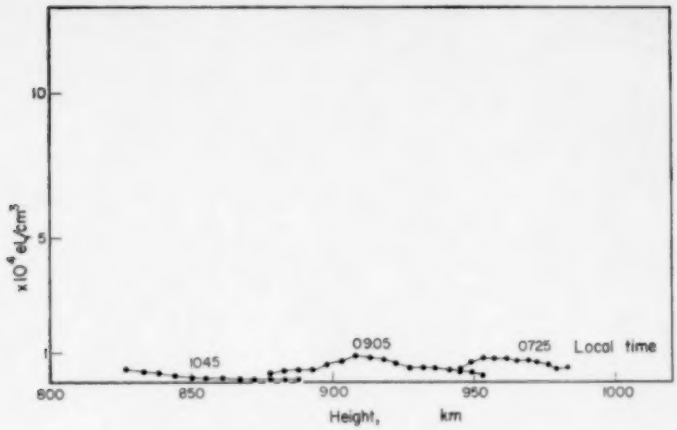


Fig. 5. Electron density profiles, 29th Sept. 1959.

The electron-density profits calculated by means of formula (26) are shown in Figs. 4 and 5. It will be seen that positive gradients, corresponding to differences in electron density by a factor 2–4, may occur in the ionosphere above the *F2* maximum.

Acknowledgements—The author wishes to thank the Director of Kiruna Geophysical Observatory, Dr. Bengt Hultqvist, for invaluable discussions. A number of discussions with Professor H. Alfvén and Dr. J. H. Thomson have also been most helpful. The work on which this communication is based has been supported by the European Office of Air Research and Development Command (Contract 61(514)-1314).

REFERENCES

1. G. J. Aitchison and K. Weeks. *J. Atmos. Terr. Phys.* **14,** 236 (1959).
2. O. K. Garriott, *J. Geophys. Res.* **65,** 1139 (1960).
3. F. B. Daniels, p. 276 in *Scientific Uses of Earth Satellites* (edited by J. A. Van Allen), University of Michigan Press, Ann Arbor (1958).
4. E. H. Vestine, L. Laporte, I. Lange and W. E. Scott. *The Geomagnetic Field, Its Description and Analysis.* Carnegie Institution of Washington Publication No. 580 (1947).

SELECTED PAPERS AND ABSTRACTS TRANSLATED FROM

ARTIFICIAL EARTH SATELLITES
U. S. S. R.

(ISKUSSTVENNYE SPUTNIKI ZEMLI)

(PUBLISHED BY THE U.S.S.R. ACADEMY OF SCIENCES, MOSCOW)

Translations of any papers abstracted here may be obtained from Pergamon Institute at a reasonable cost. Enquiries should be addressed to the Administrative Secretary, Pergamon Institute, at either Headington Hill Hall, Oxford, England, or 122 East 55th Street, New York 22, N.Y., U.S.A.

Any papers marked * have been selected for publication in this section of future issues of *Planetary and Space Science*.

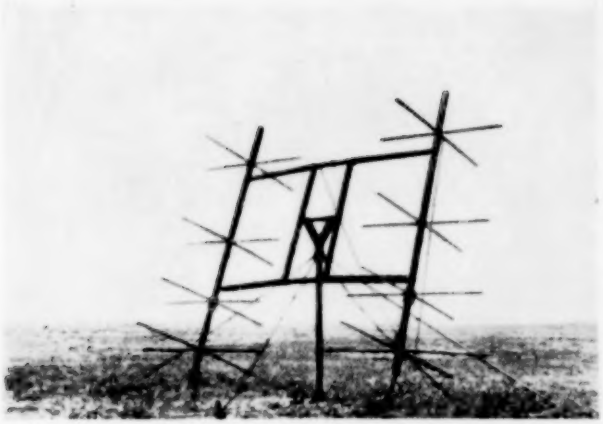


Fig. 1. The receiving antenna used in the experiment with the ultra short wave dispersion interferometer.

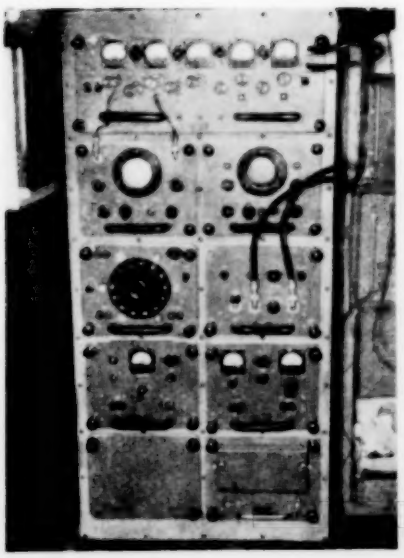


Fig. 2. The phase-meter receiver of the dispersion interferometer.

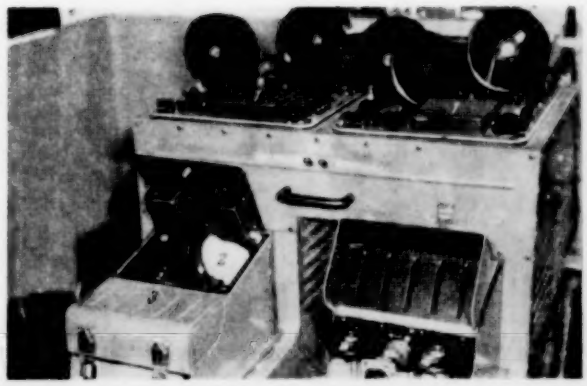


Fig. 3. The recording oscillograph of the dispersion interferometer.
1 = Loop oscillograph; 2 = tube for recording the characteristic points of the Lissajous figures; 3 = photographic recorder.

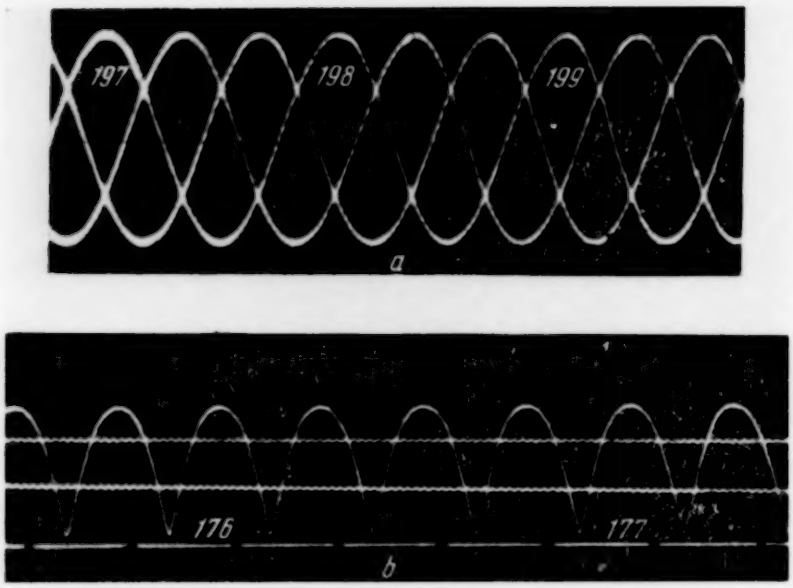


Fig. 4. Specimen records.
a = displacements of the characteristic points of the Lissajous figures; b = interference frequency.

Fig. 10. General view of the ion trap.

(*To face page* 223)

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 223-232. Printed in Great Britain.

SOVIET INVESTIGATIONS OF THE IONOSPHERE EMPLOYING ROCKETS AND ARTIFICIAL EARTH SATELLITES*

V. I. KRASOVSKII

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli* No. 2, p. 36 (1958).

The methods of investigating the ionosphere may be divided into four groups:

(1) using radio waves from ground radio transmitters reflected from the ionosphere;

(2) using radio emissions from extra-terrestrial sources passing through the ionosphere (the sun and radio sources) and from ground sources employing Moon radar;

(3) investigating radio emission of sources moving about in the ionosphere itself, employing rockets and artificial earth satellites;

(4) using instruments for directly determining the characteristics of the ionosphere in proximity to the rocket or satellites (Langmuir probes, ion traps, mass spectrometers, etc.), whose readings are transmitted to earth by radio telemeter transmission.

Until quite recently the main information regarding the ionosphere was obtained from investigations of the propagation of radio waves obtained from equipment located on the ground. Only the method of pulse radiosounding of the ionosphere was widely adopted. However, the study of signals reflected from the atmosphere could not in principle resolve the problem of electron concentration distribution over the entire vertical section of the atmosphere, since this method did not provide an exhaustive investigation of the broad layers of the inonosphere where the ionisation density falls off with height or does not exceed the ionisation density in the lower zones. In actual fact, whilst an analysis of the group delay of radio-wave impulses of different frequencies reflected from the F zone provides in some measure conclusions regarding electron concentration above the ionisation maximum of the E zone, the zone occurring above the maximum of the highest layer of those observed by means of reflected radio waves cannot generally be studied by this method.

In ionosphere investigations employing radio emission passing through the entire thickness of the earth's atmosphere from extra-terrestrial sources it is possible to determine the integral content of electrons throughout the entire vertical column. By comparing these data with observations of ionosphere stations it is possible to derive conclusions regarding the electron concentration of the ionosphere zones above the maximums of the layers. This method, however, does not provide data regarding the true electron concentration distribution according to height in all its details.

Hence exhaustive data regarding electron concentration distribution over the whole vertical section of the atmosphere cannot be obtained by means of instruments on the ground. This problem can be solved by means of radio methods only if radio wave sources, operating on frequencies exceeding the critical frequencies of the ionosphere layers, are arranged vertically throughout the ionosphere itself. It is important in this respect that the vertical displacements of the emitting sources should substantially exceed their horizontal displacements. This condition is well achieved in the vertical launching of high-altitude rockets. A wealth of possibilities are opened up for the investigation of ion and electron concentrations in the atmosphere by employing direct measurements on rockets and in particular on artificial earth satellites.

A number of works have been published in recent years setting out measurements of electron concentration in the ionosphere in the U.S.A. employing rockets.[1-7] The majority[1-5, 7] relate to zones up to 200 km and only one extends to an altitude of 380 km.[6]

* The present paper is a review of the main works on ionosphere research carried out in the U.S.S.R. by means of rockets and satellites.

In the Soviet Union prime importance attaches to ionosphere studies by means of rockets and artificial earth satellites. We will deal first of all with ionosphere investigations based on the study of radio wave propagation. The most important investigations in this field take the form of measurements of electron concentrations in the ionosphere by the dispersion interferometer method, carried out from 1954 at the time of the launchings of the geophysical rockets by the Soviet Academy of Sciences.[6] The dispersion interferometer method is one of the radio interference methods for studying radio wave propagation proposed by Mandel'shtam and Papaleksi more than twenty years ago. This method was already being used in 1936 for the study of ionosphere variations of the optical path of radio waves with wavelengths in the order of some hundreds of metres,[9-11] the radio transmitting and receiving sections of the interferometer being installed at different points on the earth's surface.

During the time of the flights of the Soviet artificial earth satellites mass measurements were taken of signal levels of the radio waves received from the satellite. Treatment of the results of these measurements provided interesting conclusions regarding the characteristics and state of the ionosphere.

Al'pert *et al.*[12] worked on these measurements in order to establish the possibility of using the phenomena of "radio ascent" and "radio descent" of the satellite for obtaining data regarding the structure of the ionosphere. Kazantsev and others[13] treated the measurement results from the standpoint of determining radio wave absorption above the maximum of the *F2* layer. We will touch on each of these works.

On the basis of the determination of electron concentration, dispersion measurements were taken of the radio waves emitted from the rocket by the dispersion interferometer method. The experiments were carried out in the following manner. On a rocket, launched at a small angle to the vertical, radio transmitters were installed, transmitting coherent radio waves in the ultra-short wave range on two frequencies, the one being a multiple of the other. This was achieved by discriminatory frequency multiplication of the common quartz-stabilised generator. During flight of the rocket the radio waves from these two transmitters were received at two points on the ground, and the differences in phase and intensity of the received oscillations were recorded continuously. Measurements of the rocket co-ordinates were carried out simultaneously both by optical and radio methods together with radio sounding of the ionosphere by means of a panoramic ionosphere station near to the rocket launching site. For the purpose of increasing the phase sensitivity of the phase-meter receiving equipment, use was made of sequential frequency multiplication and heterodyning of the received signals. Phase-difference recordings were carried out by two methods: continuous photography of the characteristic points of the Lissajous figures from the cathode ray oscillograph tube by the method described by Shchegolev *et al.*[14], in addition to the method of recording on the loop oscillograph the interference frequency formed after detecting the sum of the two signals received. In the latter case the signal with the longer frequency was initially multiplied in frequency by the number of times that this frequency was below the frequency of the high frequency signal. Photographs of the equipment and specimen records are shown in Figs. 1–4.

Appendix 1 provides the final formula for calculating the mean electron density at a given altitude, showing the result and the tolerances adopted.

Launchings of these Soviet Academy of Sciences geophysical rockets were carried out in the middle latitudes of the European U.S.S.R. In constructing the electron density distribution curves as a function of altitude use was made of data obtained only on the ascending sections of the rocket trajectories (*i.e.* while ascending).

During launchings in the 1954–1957 period frequencies of 48 and 144 Mc/s were employed (with triple multiplication of the lower frequency). At the time of the experiment on 21st February 1958 additional records were also made of the phase difference of radio

waves employed with frequency of 144 Mc/s and the coherent radio waves with frequency of 24 Mc/s, also emitted from the rocket. At the time of this launching a record was made of the rotation of the polarisation plane, for all the radio waves received, due to magneto-ionic fission in the ionosphere. Ultimate treatment of the 1958 material has not yet been completed and the data set out are averaged over approximately 10-km altitude intervals. It is an important fact that the geophysical rocket, having risen to an altitude of 475 km, after switching off of the engine, was stabilised relative to three possible axes of rotation.

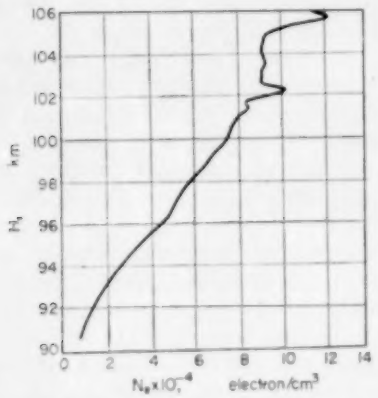


Fig. 5. Electron concentration distribution against altitude.
17 h 24 m 26th May 1954.

The results of a number of experiments are illustrated in Figs. 5–7. The captions to the diagrams show local time. The 1957–1958 data point to the conclusion that at the investigated altitudes, in addition to the *F* zone, there are no sharply defined layers in the ionosphere; the ionosphere consists of a very great number of small electron densifications against a general background of a monotonous rise in electron concentration to the maximum of the *F2* layer. This all agrees with the American data.[1–7] However, the results of measurements, carried out on 21st February 1958 above the maximum ionisation of the *F2* layer, differ considerably from Berning's data,[6] according to which the electron concentration rapidly falls off above the maximum of the layer and is very small already at an altitude of 378 km. In our experiments it varies from a value of $1{\cdot}8 \times 10^6$ electrons/cm^3 at an altitude of 290 km to a value of 10^6 electrons/cm^3 at an altitude of 473 km. It should be observed that at the time

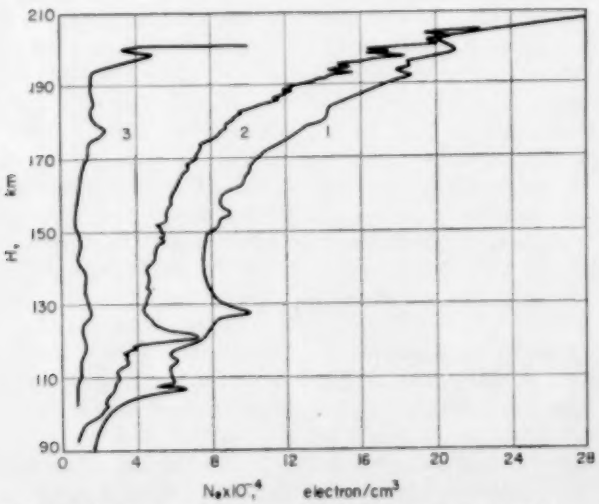


Fig. 6. Electron concentration distribution against altitude.
1 = 6h 18m 16th May 1957; 2 = 6h 27m 25th Aug. 1957; 3 = 19h 54m 9th Sept. 1957.

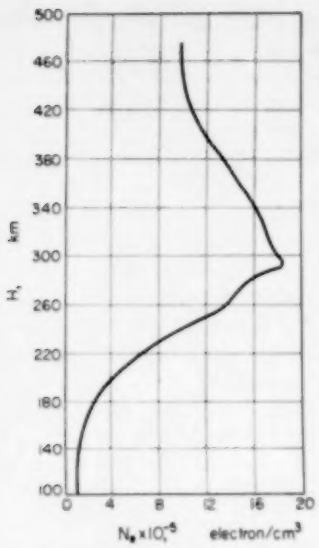


Fig. 7. Electron concentration distribution against altitude, 11^h 40^m 21st Feb. 1958.

of these experiments the phases of the solar activity cycle, time of year and day were completely comparable in the U.S.A. and the U.S.S.R.

The 1957–1958 results were compared with simultaneous measurements at ionosphere stations located near to the launching site of the rocket. It was found that the altitudes recorded by the ionosphere stations in the *F* zone exceed by 50–150 km the altitudes obtained by means of the dispersion interferometer.

During the life of the Soviet artificial earth satellites a wide range of material was accumulated on the reception of satellite signals and the intensity level of the signals. This material is still in the treatment stage. Individual selected investigations are set out at the present stage.

An investigation was made of the signal levels from the radio transmitters of satellites at 20·005 Mc/s. By comparing the mean field strengths of actual signals with the field strengths as calculated from the formula for ideal radio transmission, radio wave absorptions were determined for different positions of the satellite. Comparison of the radio wave absorption when the satellite was above the ionisation maximum and below the *F2* layer maximum provided an assessment of the density variation of the ionosphere above the *F2* layer maximum. According to these evaluations, reduction in the electron concentration above the layer maximum occurs considerably more slowly than increase when approaching the minimum from below.

It was observed in a number of cases that the radio waves approached the point of reception not by the shortest distance but by passing around the earth along the longer arc of a large circle. In some cases the measured field strength values were greater than the calculated, following a law in inverse proportion to the first power of the distance, which evidently points to the propagation of radio waves in these cases along some particular channel in the ionosphere.

An investigation was next carried out into the phenomena of radio ascent and radio descent when receiving radio signals from satellites on a frequency of 40 Mc/s. In some tens of cases taken from several hundred it was possible to observe "radio ascent" and "radio descent" of the satellites in pure form and to establish the corresponding moments in time. Radio emission is subjected to very intense distortion in the ionosphere. The lag in true ascent as compared with radio ascent and delay in radio descent as compared with true descent permit a determination of the amount of distortion of the radio emission. Since distortion of the radio emission in the ionosphere depends on the variation in electron concentration with altitude then it is possible—given the different laws of electron concentration variation to some degree of discrepancy—to select some electron concentration distribution with altitude which would yield the radio ascent and radio descent times actually observed. The influence of the lower layers of the ionosphere may be calculated to a certain degree on the basis of direct measurements carried out by the network of ground stations.

Data obtained as a result of observations of radio signals of the first artificial earth satellites lead to the conclusion that the electron concentration value in the outer ionosphere

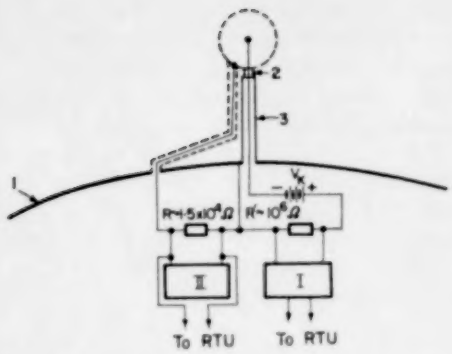


Fig. 8. Experimental layout using ion traps. 1=Surface of the satellite; 2=insulator; 3=adjustable screening tube.

falls with altitude much more slowly than it increases below the maximum. Hence there is a qualitative agreement between these evaluations of electron content in the ionosphere with those obtained by means of rocket dispersion radio interferometers.

In addition to the use of radio waves emitted from the satellite for studying the ionosphere, an instrument was installed on the third Soviet earth satellite for the direct measurement of charged-particle concentration in the ionosphere, with the transmission of the instrument readings on the radio telemetering system. A feature of this type of experiment is the independence of the measurement results from the characteristics of the overall thickness of the ionosphere between the earth and the satellite and from the processes occurring therein.

The experiment carried out on the third Soviet artificial earth satellite was proposed by Gringauz and Zelikman.[15]

The ionosphere is quasineutral, *i.e.* it can be considered that the positive-ion concentration is equal to the sum of the concentrations of electrons and negative ions. Hence the total charged-particle concentration is equal to twice the positive-particle concentration. Measurement of the latter affords determination of the former.

The apparatus installed on the third Soviet satellite for measuring positive ions consists of two spherical wire-gauze ion traps (Figs. 8–10), fitted to thin rods and set up on diametrically opposite sides of the satellite surface (Fig. 9), together with an electronic unit consisting of two amplifier-converters and a saw-tooth impulse voltage generator (Fig. 8). Each ion trap takes the form of a wire-gauze sphere with diameter 10 cm (Fig. 10). A spherical collector is fitted in the centre, to which a negative voltage is applied relative to the wire-gauze casing (~−150 V). The electric field inside the trap accumulates on the collector all the positive particles reaching the trap and repels all the negative particles.

The collector is connected to the conducting surface of the satellite through a resistance (Fig. 8). The flow of positive ions attracted to the collector flows off to the surface of the

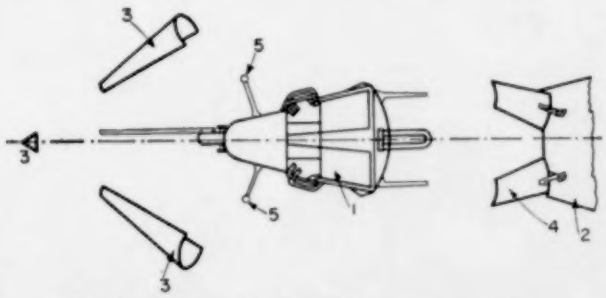


Fig. 9. Diagram showing position of traps on the satellite (moment of separation of the satellite from the rocket-carrier).
1=Satellite; 2=rocket carrier; 3=disengaging protective cone; 4=shields disengaged from satellite; 5=ion traps.

satellite, giving rise to a corresponding growth in the electron flow from the surrounding medium to the surface of the satellite. The voltage from the input resistance is applied to the amplifier input, the output of which is connected to the radio telemetering system, transmitting the instrument readings to earth.

The concentration of positive ions can be determined from the formula set out in Appendix 2.

From various causes (difference in the speeds of electrons and ions, photo-emission of electrons from the satellite surface, etc.) the satellite may take on an electrical charge, as a result of which the potential of the wire-gauze shell of the trap relative to the ambient medium will vary from zero. In order to take account of the influence of the electrical charge of the satellite on the measurement results, use is made of saw-teeth impulse voltages, applied to the wire-gauze shell relative to the body of the satellite.

At the moments of application of the saw-tooth impulses, simultaneous telemeter recording of the collector-trap currents and the voltage between the surfaces of the traps and the satellite surface provides volt–ampere characteristics of the current at the collector as a function of the potential difference between the wire-gauze sphere of the traps and the body of the satellite. The anticipated character of the characteristics is shown in Fig. 11.[15] The point *A*, where the collector current ceases to diminish with increasing potential of the wire-gauze sphere of the trap relative to the surrounding medium, corresponds to the ion retardation potential, which can be determined on the basis of the familiar expression set out in Appendix 3. Knowing the impulse generator voltage and the ion retardation potential the point can be found on the curve of Fig. 11 corresponding to the condition for which the potential of the wire-gauze sphere of the traps relative to the surrounding medium will be equal to zero. At this time the ion flux density at the trap reflects the true density in the ionosphere. This affords a determination of the true ion flow in the intervals of time between two adjacent saw-tooth impulses. The current corresponding to the retardation potential is the total current formed by parastic processes.

The record of collector currents in the absence of voltage impulses at the traps affords a measurement of the dimensions of small-scale ionisation disuniformities.

At present the treatment of the results of this experiment is only in the initial stage, and it is therefore possible only to provide preliminary information obtained from the performers of the experiment, characterizing the measurements taken. We will limit ourselves to a study of measurements at two points on the satellite orbit, relating to the first day of the satellite flight. These results are typical, although other values have been observed.

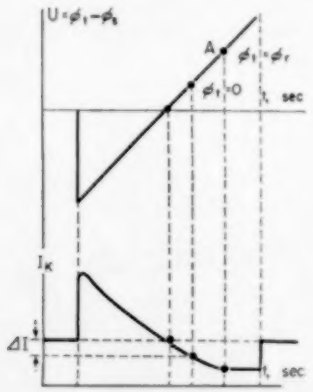


Fig. 11. Shape of the saw-tooth voltage impulses (upper curve), applied to the wire-gauze sphere of the trap, and shape of anticipated signals (lower curve).

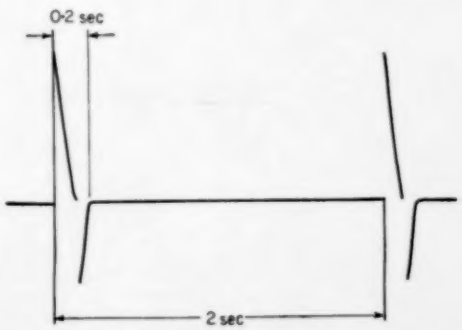


Fig. 12. Application of saw-tooth impulses over a time interval.

Every 2 sec, bipolar voltage impulses are applied to the wire-gauze sphere of the ion traps relative to the body of the satellite, with an overall duration of 0·2 sec (Fig. 12).

Figure 13 shows the volt–ampere characteristics of the collector currents of the traps as a function of voltage on the spherical spheres of the traps at an altitude of 795 km, 15th May 1958, 5 to 6 hr after midday at the middle latitudes. The nature of the curves is as expected (see Fig. 11).[10] One of the ion traps at the given moment is in the rarefied zone formed behind the satellite, as the result of which the ion flow in it is very small. There can be observed on the curves points of complete positive-ion retardation. Considering that the heaviest ion M_{max} at this altitude is the atmospheric oxygen ion (mass number 16) and that consequently the kinetic energy of the ions at the moment of complete retardation $E_T=(M_{max}\cdot v^2 s)/2e\sim 5{\cdot}1$ e V, it can be found from the characteristics (Fig. 13) that the potential of the sphere of the trap is equal to the potential of the surrounding plasma on application of a voltage of $+6{\cdot}4$ V relative to the body and that, consequently the negative potential of the body of the satellite relative to the plasma is equal to 6·4 V. Such a satellite potential (even disregarding the influence of photoeffect) can occur only with an electron effective temperature in the order of tens of thousands of degrees Kelvin, that is to say at such a temperature as would in the case of a Maxwellian distribution of electron speeds lead to a negative potential of the body as indicated, disregarding the influence of photoemission from the surface. The positive-ion concentration, determined from the current, corresponding to zero potential of the sphere of the trap relative to the plasma, is equal to $1{\cdot}9\times 10^5$ ion/cm^3.

On applying voltage impulses to the wire-gauze spheres of the traps these appear as parallel-connected spherical Langmuir probes. On applying a positive voltage relative to the body of the satellite the electron current at the probes increases sharply, giving rise to a drop in voltage at the output of the saw-tooth impulse generator. Knowing the difference in shape of the positive voltage impulse from its undistorted shape it is possible to construct the probe characteristic $\log I_3=f(V)$, where I_3 is the total current at both spheres of the traps, thence determining the current at which the satellite potential is equal to the plasma potential, together with the electron temperaure.

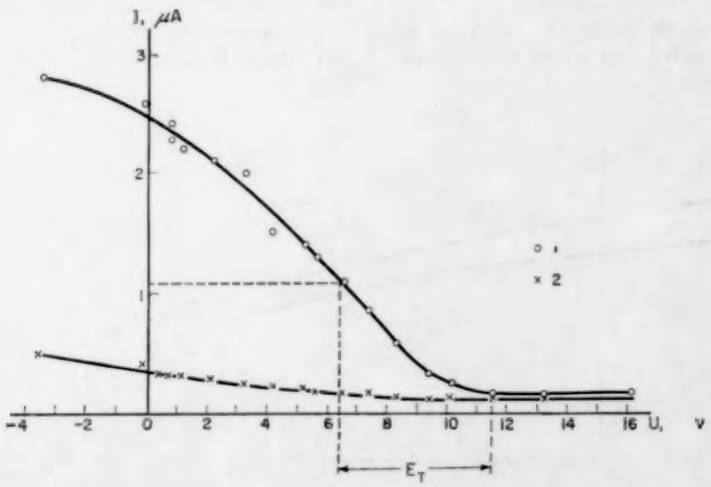


Fig. 13. Specimen curves showing currents at the collectors of ion traps as a function of potential of wire-gauze sphere.
1 = current at the first collector; 2 = current at the second collector.

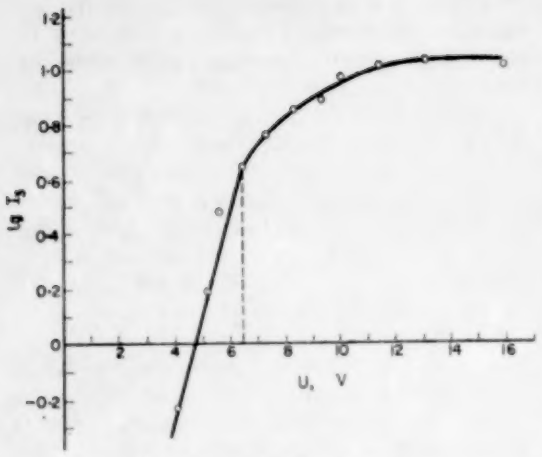


Fig. 14. Probe characteristic.

Figure 14 shows the total current at both wire-gauze spheres of the trap as a function of the potential (probe characteristic) for the same experiment as in Fig. 13 ($H = 795$km, 15th May 1958). The potential determined from this characteristic corresponds to the plasma potential, approximating to that determined from the volt–ampere characteristics of the ion currents, the electron temperature being not less than 15000°K.

In several cases (at lower altitudes) the positive voltage impulse diminishes as a result of the drop in voltage caused by the electron currents at the wire-gauze sphere of the traps to such a degree that total retardation of the positive ions does not occur. By way of example we can show the volt-ampere characteristics obtained at an altitude of 242 km on 15th May 1958, 1–2 hours before midday at the middle latitudes (Fig. 15). Under these conditions, in determining the point at which the potential of the wire-gauze sphere is equal to the plasma potential, only the probe characteristics should be employed. The two curves represent the currents at different collectors. In the given case, at an altitude of 242 km the positive-ion concentration was equal to $5{\cdot}2 \times 10^5$ ion/cm^3. the negative satellite potential$\sim$2·0 V and the electron effective temperature $\sim$7000°K.

Determinations of the ion composition of the upper atmosphere were also carried out on rockets and on the third Soviet artificial earth satellite.[16]

CONCLUSIONS

(1) Observations of radio waves emitted from artificial earth satellites carried out for various satellite positions relative to maximum ionisation of the *F4* layer can provide the basis for valuable information regarding the characteristics and state of the ionosphere. The observations should be extended in future, the study of simultaneous passage of radio waves at

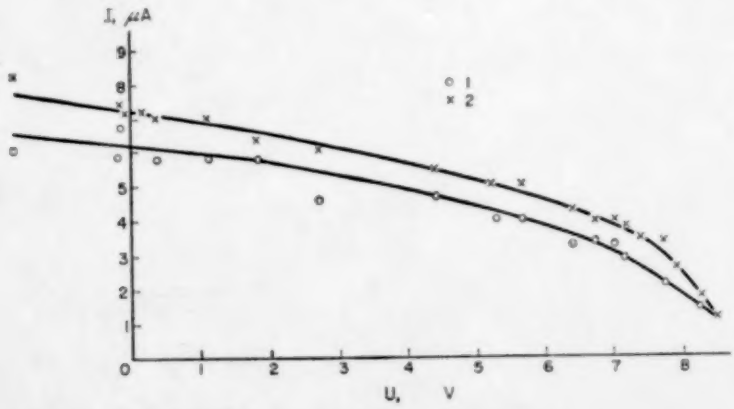


Fig. 15. Ion current as a function of retardation voltage.
1 = current of the first collector; 2 = current of the second collector.

different frequencies through the ionosphere being particularly necessary.

(2) Other investigations of the ionosphere by means of rockets and artificial earth satellites carried out in the U.S.S.R. have up to the present provided the following results:

(a) Ionisation above the maximum of the *F2* layer falls off very slowly with altitude, pointing to a very large scale of altitudes for ionised particles in the ionosphere; on 21st February 1958 ionisation of approximately 10^6 electrons/cm^3 was observed at an altitude of 473 km; on 15th May 1958 a positive-ion concentration amounting to $1{\cdot}9\times10^5$ ions/cm^3 was measured at an altitude of 795 km.

(b) The effective electron temperature in the *F* zone and higher is several times greater than the temperature of the medium.

(c) It has been confirmed that the ionosphere below the maximum of the *F* zone, as distinct from the model, constructed from ionosphere sounding data from the earth, does not possess sharply defined layers and is characterized by a number of small maximums with monotonous increase of the electron density upwards towards the maximum of the *F2* layer.

(d) It has been confirmed that the maximum of the *F2* layer and generally ionisation below the *F* zone is 50–150 km below the values emerging from the data of usual ionosphere sounding from the ground.

The material obtained is very important from the standpoint of the physics of the upper atmosphere. The large scale of altitudes for electrons above the maximum of the *F2* layer in addition to the high effect of electron temperature is in close harmony with the concept of intense heating of the upper atmosphere over a large scale of altitudes, which emerges from observations of retardation of artificial earth satellites. The absence of thermodynamic equilibrium between the electrons, ions and other particles of the upper atmosphere is a very important feature. It is quite possible that the electrons of the upper atmosphere are accelerated in short-period alternating geomagnetic fields, set up by circulation of the atmosphere or of the interplanetary medium, giving rise to heating of the upper atmosphere.

The multi- and fine-layer structure of the ionosphere below the maximum of the *F2* layer, together with its lower location than from direct data of radio sounding from the ground, demands serious revision and improvement of the old model. Evidently ordinary ionosphere sounding provides a coarser effective representation of the ionosphere than investigation by means of rockets and artificial earth satellites. It is of course obvious that this does not detract from the practical value of ionosphere sounding, which has been and will continue in the future to be a practical means for obtaining effective ionosphere parameters. However, an exhaustive and thorough investigation of the ionosphere is possible only by means of rockets and artificial earth satellites. Striking advances have already been achieved in this direction. Fascinating prospects lie ahead in the investigation of the upper atmosphere at ever-increasing distances from the earth's surface.

APPENDIX I

Assume

$$n_f=\sqrt{\left(1-\frac{N_e^2}{\mu mf^2}\right)}\approx 1-\frac{N_e^2}{2\mu mf^2}, \quad . \quad (1)$$

where n_f is the refractive index in the ionosphere for the frequency f; N is the effective concentration in electrons/cm^3 and e and m the charge and mass of the electron respectively.

The phase difference Φ at the point of reception of two radio waves by reason of their dispersion (referred to the higher frequency) can be represented as

$$\Phi=\frac{2\mu pf}{c}\int_0^L (n_{pf}-n_f)\,dL, \quad . \quad . \quad (2)$$

where L is the distance from the transmitter to the receiver; p is the ratio of frequencies of the radio wave on the higher frequency to the radio wave on the lower frequency; c is the speed of light cm/sec. The variation in $\Delta\Phi$ for a change in L from L to $L+\Delta\Phi$, on the basis of equation (1), can be set out as

$$\Delta\Phi=\frac{e^2}{mcf}\left(\frac{p^2-1}{p}\right)\int_L^{L+\Delta L} N\,dL \quad . \quad . \quad (3)$$

and the mean value of N along the path ΔL

$$N = \frac{1}{\Delta L}\int_{L}^{L+\Delta L} N\,dL = \frac{cmf}{c^2} \cdot \frac{p}{p^2-1} \cdot \frac{\Delta\Phi}{\Delta L} \quad . \qquad (4)$$

It is assumed that during the time of variation of L by the value ΔL the value $\int_0^L N\,dL$ does not vary.

The difference in ΔL from Δh—increase in altitude above the ground—is calculated by treating the measurement results.

APPENDIX II

If the ion trap is in a stream of ionised gas then the measured collector current is

$$I_+ = N_+ e v_s \mu r^2 \pi f(\varphi_t),$$

where e is the electron charge; r is the radius of the wire-gauze surface of the trap; α is its transmittence factor; v_s is the speed of the satellite; φ_t is the potential of the wire-gauze sphere of the trap relative to the surrounding medium.

APPENDIX III

The retardation potential φ_T can be determined from the relationship

$$e\varphi_T = \frac{M_{i\max} v_s^2}{2}$$

where e is the electron charge; $M_{i\max}$ is the mass of the heaviest ions entering the trap; v_s is the speed of the satellite.

REFERENCES

1. J. C. Seddon, *J. Geophys. Res.* **58,** 323 (1953).
2. J. E. Jackson, *J. Geophys. Res.* **59,** 577 (1954).
3. J. C. Seddon, *J. Geophys. Res.* **59,** 463 (1954).
4. J. C. Seddon, A. D. Pickar and J. E. Jackson, *J. Geophys. Res.* **59,** 513 (1954).
5. J. R. Lien, R. J. Marcon, J. C. Venrick, J. Aarons and M. Morrow, *Rocket Exploration of the Upper Atmosphere* (Eds. R. L. F. Boyd and M. J. Seaton), Pergamon Press, London (1954).
6. W. W. Berning, *J. Met., N.Y.* **8,** 175 (1951).
7. C. Hok and W. G. Dow, *Rocket Exploration of of the Upper Atmosphere* (Eds. R. L. F. Boyd and M. J. Seaton), Pergamon Press, London (1954).
8. K. I. Gringaux, *Dokl. Akad. Nauk SSSR* **120,** 1934 (1958).
9. N. D. Papaleksi, *Trudy Ekspeditsii po Nablyudeniyu Solnechnogo Zatmeniya, 19 Iyunya 1936. (Transactions of the Expedition for Observing the Solar Eclipse, 19th June 1936.) Akad. Nauk. SSSR* p. 115 (1937).
10. L. I. Mandel'shtam and M. D. Papaleksi (Editors), *Noveishiye Issledovaniya Rasprostraneniya Radiovoln vdol' Zemnoi Poverkhnosti (Recent Investigations of Radiowave Propagation Along the Earth's Surface).* Gostekhizdat, Moscow (1945).
11. E. Ya Shchegolev, *Nauka i zhizn'* No. 4 (1937).
12. Ya. L. Al'pert, E. F. Chudesenko and B. S. Shapiro, *U.F.N.* **65,** 161 (1958).
13. A. N. Mazantsev, G. S. Romanova and A. Ya. Klementenko, in press.
14. E. Ya. Shchegolev, J. E. Viller and I. M. Borushko, *Noveishiye Issledovaniya Rasprostraneniya Radiovoln vdol' Zemnoi Poverkhnosti (Recent Investigations of Radiowave propagation along the Earth's Surface).* Gostekhizdat, Moscow (1945).
15. K. I. Gringauz and M.Kh. Zelikman, *U.F.N.* **63,** 239 (1957).
16. V. G. Istomin, this issue, p. 252 (Abstract).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 233-237. Printed in Great Britain.

OPTICAL METHODS FOR TRACKING EARTH SATELLITES

I. S. SHKLOVSKII

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli* (Artificial Earth Satellites) No. 1, p. 44 (1958)

In connection with launching artificial earth satellites there is the important problem of determining their space co-ordinates for various moments of time by means of observations from the earth's surface. For determining the satellite co-ordinates use is made both of radio and optical methods; the present paper is confined to optical methods. In principle the optical methods allow greater precision in the determination of the satellite co-ordinates than the radio methods. In addition, optical observations provide the true angular co-ordinates of the earth satellite whilst the results of radio measurements are distorted in varied degrees through the influence of the ionosphere.[1]

A knowledge of the precise satellite co-ordinates is necessary for calculating its orbit. Important data can be obtained on the basis of such measurements relating to the physical characteristics of the upper layers of the earth's atmosphere. Sufficiently accurate satellite orbit parameters (such as can be obtained only from high-quality observations) render possible the determination of the density distribution of the upper layers of the earth's atmosphere. In addition it is possible to study tide effects in the upper layers of the atmosphere, which apparently attain a high value. It is very important to be able to discern characteristic deviations in the density distribution of the upper atmosphere from spherical symmetry.

The analysis of the satellite path is very important in regard to the problem of the earth's form. This question has attracted the attention of astronomers and geodesists for centuries. Its solution demands very accurate determinations of time and stellar co-ordinates. Artificial earth satellite observations offer a solution to this problem even if determination of the co-ordinates is fairly coarse (stellar co-ordinates are measured at present with an accuracy of some hundredths of an angular second). Observation of the flight of an artificial earth satellite within an accuracy of a few seconds of the arc (with corresponding accuracy in recording the moment in time) permits determination of the co-ordinates of the point of observation with an accuracy of a few metres (understanding that the satellite orbit is known).

To an observer on the earth's surface an artificial earth satellite (if illuminated by the sun) is a fairly faint, practically point object, travelling across the firmament at an appreciable angular velocity. The stellar magnitude of an artificial earth satellite can be readily calculated from a knowledge of its effective reflecting surface, its reflectance coefficient and distance to the satellite. Hence, for an artificial earth satellite with a 50-cm diameter and with mirror reflection, calculations yield stellar magnitudes as shown in Table 1.[2]

Table 1

Distance from observer (km)	Photovisual stellar magnitude (m)	Angular velocity (ft/sec)
360	5·7	88
720	7·2	42
1080	8·1	27
1440	8·7	19
1800	9·2	15
2160	9·6	12
2700	10·1	9

Observations of artificial earth satellites have shown that theoretical evaluations are in close agreement with observed values of the stellar magnitude of a satellite.

The artificial earth satellite co-ordinates can be determined by photographing the image of the satellite and the surrounding stars.

The efficacy of the photographic telescope

for recording a moving point object is determined by the parameter

$$SD^2/\omega dF,$$

where S is the sensitivity of the photo emulsion; ω is the angular velocity of the object; d is the diffusion circle of the photo emulsion; D is the diameter of the objective and F is the focal length of the objective.

Using a sufficiently high-speed system of large diameter in conjunction with sensitive radiation receivers (photographic plate, image converter tube, television system) it is possible to observe very faint artificial earth satellites.

However, the achievement of a sufficiently high sensitivity is only one aspect of the problem; artificial earth satellites have nearly always to be observed against the background of twilight illumination; sometimes bright artificial earth satellites can be observed with the naked eye when not a single star is visible in the sky. Hence an important part is also played by the contrast sensitivity of the equipment. In the case of visual observations of artificial earth satellites the contrast sensitivity of the eye varies sharply in relation to adaptation; this deteriorates in the case of dark adaptation.

When employing the photographic method, the background is determined by the speed of the instrument and the brightness of the sky; in practice when observation is being carried out against a twilight background it is necessary to employ fairly short exposures; however, there is a danger here of an inadequate number of stellar images on the photographic plate from which to determine the artificial earth satellite co-ordinates. These factors also limit the sensitivity of the method when employing image converter tubes.

The importance of high resolving power of the equipment must be emphasised; in the case of a low resolving power the contrast "artificial earth satellite trace–background" falls off. A higher resolving power is provided by the Schmidt systems; their optical speed however must not be too high since otherwise observations of artificial earth satellites in twilight are rendered difficult. The accuracy of measurement of artificial earth satellite co-ordinates depends on the accuracy of recording the moment of exposure. In the event of insufficient speed of actuation of the shutter the end of the satellite trace on the photographic plate is fogged, making it difficult to apply the filaments of the measuring instrument over it. The geographical co-ordinates of the camera site must be known very accurately.

Apart from the small spherical satellite, observation of which is fairly difficult on account of its feeble brightness, the observation of larger and brighter artificial earth satellites is also important. These can be tracked by simpler photographic equipment employing a smaller-diameter objective. Equipment for photographing bright artificial earth satellites has been set up in the Shternberg State Astronomical Institute. The photograph is taken on photographic plates (for improving the accuracy of artificial earth satellite co-ordinate measurements). Some difficulties arose in establishing a precise record of the moments of opening and closing of the camera shutters. Mechanical contacts fitted on the moving parts of the shutter require investigation with regard to the delay in the moment of contact-closing relative to opening and closing of the shutter; direct recording of the opening and closing moments of the shutter was therefore adopted. For this purpose alongside the photographic plate, inside the camera (on the inoperative part of the camera field of vision) there was fitted a lead sulphide photoconductive cell, illuminated through the objective and shutter by an infrared source, which did not influence the photographic plate. The current through the photoconductive cell, together with the second impulses from astronomical clocks, was recorded on a cathode ray oscillograph. Opening and closing of the shutter occur with sufficient speed (2–3 msec), to provide relatively accurate determinations of the co-ordinates of bright artificial earth satellites on this camera. The limiting stellar magnitude of an object travelling at a speed of 1°/sec for the camera described is $2m/5{\cdot}3m$ (Ref. 3).

Confusion should be avoided between accuracy in recording the moments of actuation

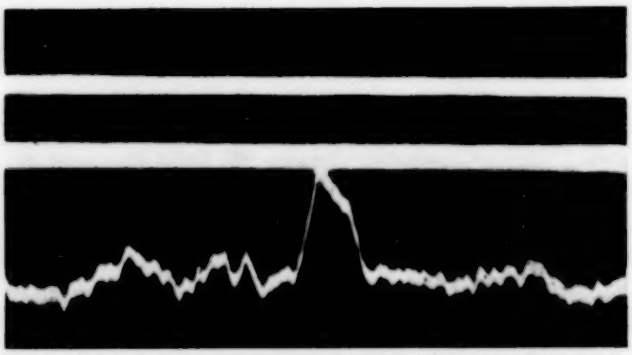


Fig. 1. Record of the passage of the second Soviet artificial earth satellite using photoelectric apparatus.

Fig. 2. Photograph of the flight of the second Soviet artificial earth satellite, obtained with an image-converter tube; brightness of the object approximately O_m, reduced 50 times.

(*To face page* 235)

of the shutter and the accuracy in determining the absolute moment of satellite flight through a given point in the firmament. As seen above, the accuracy of recording the moments of shutter actuation is approximately 2 msec, whilst correction of astronomical clocks with reference to absolute time cannot readily be achieved with such accuracy and sometimes cannot be achieved at all. However, when using for all the observations a unit time system of any given time service the recording accuracy of moments in this system will be governed only by the construction of the camera. The system itself may even differ appreciably from absolute time, but all the orbit calculations must be carried out in this special system.

The accuracy of artificial earth satellite co-ordinate measurements using the camera described above is 5 in. Calculations connected with developing the photograph are carried out on an electronic computer.

It is possible to determine artificial earth satellite co-ordinates otherwise than by photography. Imagine a slot located in the focal plane of the objective. A photoelectric cell is placed behind the slot. Let the artificial earth satellite image, moving in the focal plane, reach the slot and intersect it. At the moment of intersection of the slot the light flux falls off at the cathode of the photoelectric cell resulting in an impulse in the photoelectric current, and by recording this together with the impulses from astronomical clocks, it is possible to determine the moments of intersection of the slot edges by the artificial earth satellite image with adequate accuracy.

This type of system for noting the moment appears very attractive on account of the complete absence of any type of moving mechanical parts. The artificial earth satellite itself makes its own mark at the moment of traversing the slot (Fig. 1).

It is necessary to know the position of the slot with reference to stellar images in the focal plane. This connection can be established photographically. In the simplest case the slot can be cut in the emulsion of the photographic plate and then the stars will be photographed on the plate. If the artificial earth satellite is bright, as in the case of the rocket-carrier of the first Soviet artificial earth satellite or the second Soviet artificial earth satellite, then the trajectory of the artificial earth satellite will be obtained on this plate and the problem will be completely solved. Intersection of the trajectory with the slot provides a point for which the moment in time is known.

In the case of a faint artificial earth satellite which is incapable of providing an observable trace on a photographic plate, two systems of differently orientated slots are necessary, located in relationship to the stars. For this purpose two coupled cameras may be employed having the slots mutually perpendicular or one camera having two systems of intersecting slots, but in order to differentiate the impulses from these they must be different in width. It should be emphasised that an observable photoelectric impulse can be obtained with the passage of an artificial earth satellite for much fainter artificial earth satellites than can be recorded photographically for the same diameter aperture. This is due to the fact that the efficiency of the antimony–caesium photocathode is at least twice that of the efficiency of the photographic emulsion,* and the difference is in no way compensated by the difference in integration time in the case of artificial earth satellite observations if the time constant in photoelectric recording is equal to the efficiency of exposure. These concepts have been confirmed experimentally on actual equipment for observing artificial earth satellites.

The sensitivity of this type of photoelectric apparatus is such that in the case of deep twilights it is possible to record artificial earth satellites with a stellar magnitude of $4m/5{\cdot}5m$, when the objective diameter of the apparatus is 50 mm. When tracking in twilight conditions the sensitivity is less; it is restricted by fluctuations in the photo-multiplier current.

* The quantum yield of modern high-sensitivity plates (i.e. the number of grains of emulsion which are developed with the action of one light quantum) does not exceed 10^{-4}; the quantum yield of modern antimony–caesium photocathodes is 1/3 for the better specimens.

In designing this type of apparatus a limiting factor is the relatively small size of the photo cathode of the photomultiplier; this applies equally also to equipment using an image-converter tube. It is usually necessary to compromise between speed of the objective and the effective field, since a small field for the apparatus renders work difficult with an insufficiently accurate ephemeris. A higher optical speed increases the background influence of sky illumination and twilight.

Determination of the co-ordinates of faint artificial earth satellites can also be carried out with optical instruments having a fairly small diameter if sufficiently sensitive radiation receivers are employed. Such receivers may take the form of image-converter tubes.

In principle, image-converter devices offer very high sensitivity; their resolving power, however, is smaller than with direct photography. At present there are both contact image-converter tubes[4, 5] and multicascade systems.[6, 7]

It can be shown that the use of such systems affords an increase in the sensitivity of photography, to approximately a hundred times that obtained with the best of the existing photographic emulsions (Fig. 2). The resolving power of modern image-converter tubes is better than 0·1 mm and can be raised still further.

Disadvantages of using image-converter tubes for solving the present problem consist of the small dimensions of the effective field, coupled with a lower resolving power than with direct photography. There is at present an image-converter apparatus with objective $D=70$ mm, $D/f=1/1{\cdot}4$ and a converter with antimony-caesium cathode. In this type of equipment various types of converter with antimony-caesium cathode can be employed, in particular contact converters.

The use of television methods opens up clear perspectives in the observation of artificial earth satellites; it is known that increased contrast and discrimination of the fixed background are possible by this means.

An interesting idea was recently expounded permitting considerable increase in the brightness of artificial earth satellites.[8] Imagine that the surface of the satellite is not smoothly spherical but is multi-sided. Each face, if sufficiently polished, can be regarded as a plane mirror. These small mirrors will reflect solar "flashes" to the earth. Under some conditions the observer will see very brief splashes of light of a fairly high intensity. No calculation details are given in the paper.[8] It can be readily calculated that if a satellite some metres in size, at a distance of 500 km from the observer, possesses flat elements with an area of 100 cm², then the stellar magnitude of the "flashes" will be approximately $3m$. The diameter of the flashes on the earth will be approximately 5 km. If the satellite performs one revolution around its axis in 1 sec, then the light impulses will have a duration of approximately 3 msec, following each other approximately after 1 sec. The recording of such bright, although brief, flashes does not present difficulty on the standard cameras described above.

Difficulties are encountered however in fixing the moments of the flashes. This is due to the fact that the moment of exposure of the individual point images of the flashes is governed not by the operation of the camera shutter but by the rotation of the artificial earth satellite. In principle the moments of the flashes can be recorded by means of a supplementary photoelectric photometer, installed in the immediate vicinity of the camera.

The influence of diffraction at the edges of the mirror faces of the satellite is negligibly small, as shown by simple calculations (the radius of the first diffraction ring will be approximately 2·5 m, which is appreciably smaller than the dimensions of the "flash").

Some danger exists in carrying out this project on account of micrometeors. In the course of time bombardment of the mirror faces on the satellite by micrometeors may damage their optical surface. Unfortunately little is known so far regarding the quantity of micrometeors. Assuming that the density of micrometeors in the upper atmosphere is the same as the density of interplanetary dust, i.e. 10^{-21} g/cm³ and the size of each micrometeor is 3×10^{-4} cm, with a specific gravity 3, then the concentration of micrometeors will be 3×10^{-12} cm^{-3}. The micrometeor stream at a velocity of 3×10^{6}

cm/sec will then be 10^{-5} cm^{-2} sec^{-1}, i.e. for each square centimetre one micrometeor will impinge every 12 days. This evidently provides conservation of the mirror over a sufficiently long period. These calculations of course are only of a general nature.

Hence optical methods of tracking artificial earth satellites offer great possibilities for precise determination of the co-ordinates. Despite the fact that optical tracking of artificial earth satellites is associated with the weather and can be carried out only under twilight conditions it offers very great perspectives of development.

REFERENCES

1. H. S. W. Massey, R. L. F. Boys, *Nature, Lond.* **181,** 78 (1958).
2. K. G. Henize, *Sky and Telescope* **16,** 108 (1957).
3. I. S. Shklovskii and P. V. Scheglov, *U.F.N.* **64,** 417 (1958).
4. *Trans. 9th meet. IAU, Cambridge, 1957,* Vol. 9, p. 693.
5. W. Baum, *Astrophys Soc. Pac. Leafl. No. 326,* p. 8 (1958).
6. *Naturf. Med. in Deutschl.* **15,** 96 (1939-1946).
7. L. A. Goncharskii, USSR Copyright No. 43462 (26th Nov. 1934).
8. *Sky and Telescope* **17,** 6 (1957).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 238-243. Printed in Great Britain.

DYNAMIC EFFECTS IN THE MOTION OF ARTIFICIAL EARTH SATELLITES

L. I. SEDOV

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli* No. 2, p. 3, 1958

The laws of motion of artificial earth satellites are at present studied experimentally by optical means and radio observations, worked out on the basis of theoretical calculations of the orbit and general statistical methods for a large number of measurements. Data regarding the law of motion of the satellite along the orbit are necessary for relating the scientific measurements to points in space, for predicting the position of the satellite, for determining its life and for solving a number of geophysical problems (such as detecting and defining anomalies in the earth's gravitational field, determining the density distribution in the earth's atmosphere, etc.). Data relating to rotation and direction of the satellite in space are very important in regard to several research problems.

Theoretical calculation of the satellite orbit and its motion around the centre of gravity is connected with the analysis and calculation of the forces acting on the satellite, governed by the earth's gravity, interaction of the travelling satellite with the atmosphere surrounding the earth and with the earth's electromagnetic field.

Analysis shows that the disturbing effects of the sun, moon and meteoric particles on the motion of artificial satellites are very small and can at present be disregarded.

Problems of satellite motion taking into account only Newtonian forces of gravity resolve themselves into the familiar study of the problem of celestial mechanics, in general conformity with the problems of motion of the moon and planets. The satellite orbit in a system of co-ordinates moving progressively in relation to the stars and the earth's centre, during the course of one circle approaches an osculating ellipse at each point of the circle studied.

The position of the osculating ellipse with focus at the earth's centre can be determined by the longitude of the ascending unit Ω, the angle of inclination of the plane of the ellipse with the equatorial plane (orbital inclination i) and the distance of the perigee from the unit ω (Fig. 1); for the parameters governing the ellipse we can take its greater semi-axis a and eccentricity e. The law of motion along the circle approximates to the Kepler law of motion for the osculating ellipse.

As a result of the earth's oblateness and gravity anomalies these parameters of the elliptical orbit are subjected to disturbances and gradually vary in the course of time. The disturbances are characterized by deviations from zero, by mean circular drifts and by some small fluctuations. It is known that the main secular effect of the earth's oblateness produces a regular variation in Ω and ω. For the first Soviet artificial satellites the angle Ω varied approximately by 3°/day from east to west, whilst the longitude of the perigee ω descended approximately by $0{\cdot}4^\circ$/day.

Under the action of Newtonian gravitational forces the possible characteristic qualitative features of motion of the satellite, regarded as a solid body, with reference to the centre of gravity, can be followed on the classical examples of the motion of the moon or earth. The construction of a quantitative theory for satellites is largely connected with the different initial conditions and with the possible difference in the moments of inertia of the satellite with reference to its central axes. The fresh material effects are connected with calculation of the reciprocal forces of the satellite with the atmosphere and the earth's electromagnetic field.

Questions relating to the study of air resist-

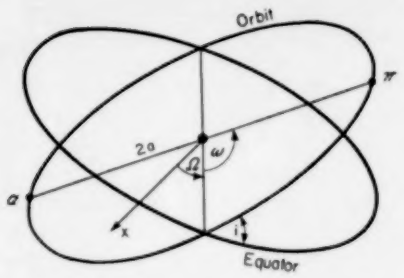


Fig. 1. Determination of the position of the orbit.

ance and its influence on satellite motion have been the subject of several works[1–5], providing a number of conclusions as to the influence of these forces, but until recently the main difficulty resided in the absence of data regarding the properties of the earth's atmosphere at high altitudes.

Since the linear speeds due to the rotation of the satellite around the centre of gravity are small as compared with the speed of the centre of gravity with reference to the atmosphere, the influence of rotation is effective only through the varying orientation of the satellite with reference to the speed vector of the centre of gravity.

In the published works the influence of aerodynamic forces was analysed on the assumption that the action of the air is reduced to the force of resistance, W, in a direction opposing the speed of the centre of gravity. Determination of the resistance W is reduced to determination of the coefficient C_x in the formula

$$W = C_x \rho \frac{S v^2}{2}, \qquad (1)$$

where ρ is the atmospheric density at that point in the satellite path, S is the total area of the outer surface, v is the speed of the satellite.

At present the coefficient C_x is determined from theoretical calculations (by the methods of molecular aerodynamics). Generally speaking, this coefficient is a function of the atmospheric temperature, satellite speed and conditions governing the interaction of molecules and atoms of the atmosphere with the particles forming the surface of the satellite, whilst at the same time the value of C_x may depend largely on the geometrical shape of the satellite and its direction in relation to the speed vector of the centre of gravity. Theoretical analysis, carried out in relation to the Soviet satellites, taking into account information relating to the rotation of the satellite around the centre of gravity, provides numerical values of C_x which may differ from the true values or vary in the course of time up to 30 per cent for the first, to 60 per cent for the second and to 30 per cent for the third satellite.

In the case of motion of a charged satellite in a plasma there is an additional resistance which, in the case of large negative charges (roughly 30 V), small satellite dimensions and at high altitude, may increase C_x to the extent of double[6]. For large-size satellites this effect is apparently immaterial. Evaluations carried out show that the force of resistance, governed by interaction of the satellite with the earth's magnetic field, is immaterial.

For altitudes of approximately 230 km the force of resistance is very small, since the atmospheric density ρ is small. According to recent data at these altitudes the density is in the order of 3×10^{-13} g/cm^3. For the third Soviet satellite the force of resistance at the perigee is in the order of $W = 4$ g and $W/G \approx 3 \times 10^{-6}$ (G being the weight of the satellite).

The published theories on satellite motion taking into account the forces of resistance have been developed on the assumption that $C_x = \text{const.}$ and that density falls off with altitude according to a certain law. In particular, use is made in practice of the following law of density diminution:

$$\rho = \rho_0 \exp - -\left(\frac{z - z_0}{H}\right), \qquad (2)$$

where ρ_0 is the density at altitude z_0, whilst z is the altitude, calculated from the normal to the surface of the earth spheroid; the value H at the isothermic sections of the atmosphere can be regarded as constant.

Observation data and calculations of the force of resistance for unorientated satellites of non-spherical shape point to the necessity for

calculating the variation in C_x as a result of rotation of the satellite relative to its centre of gravity.

Analysis shows that the presence of the resistance W leads to secular variations of the orbit, consisting of reduced eccentricity and lowering of the perigee for the osculating ellipse, accompanied by a reduced period of rotation T. The perigee descends slowly and the apogee more quickly. If the difference in height of the apogee and perigee is appreciably greater than H, then the main deceleration effect occurs in close proximity to the perigee, where the density is the greatest. It should be observed that in this case the influence of aerodynamic moments on rotation of the satellite also occurs near the perigee. It is assumed that the satellite is bounded by the surface of rotation and is dynamically symmetrical in relation to the axis of rotation. It is known that in the absence of disturbing moments relative to the centre of gravity the moment vector of momentum is constant whilst the axis of rotation of the satellite is subject to regular precession around the moment vector of momentum.

The aerodynamic forces obviously set up a small disturbance motion. The resultant of the aerodynamic forces is applied to the axis of rotation of the satellite and—other conditions being equal, for secular positions of the axis of rotation of the satellite relative to the π plane, passing through the velocity vector of the centre of gravity and the moment vector of momentum—is expressed by vectors which are mutually mirror reflections relative to the π plane. It emerges from this that the moment vector of aerodynamic forces averaged over the period of one precession revolution relative to the centre of gravity is perpendicular to the π plane. If material aerodynamic moments operate only in the region of the perigee, it follows from the preceding concepts that the secular part of disturbed rotation takes the form of a slow precession of the moment vector of momentum relative to the tangent to the orbit at the perigee. This assumption has also been established in more detail[7] by an investigation of disturbed motion.

Analysis shows that in the case of the second satellite the disturbance moments of gravity are comparable with the aerodynamic moments. It is known that the gravitational moments lead to precession of the moment vector of momentum relative to the normal to the orbital plane.

In the work cited[7] a study is also made of the simultaneous action of aerodynamic and gravitational moments, and the characteristics of disturbed motion are established taking into account regression of the orbit.

The interaction of the earth's magnetic field with the rotating metal body of the satellite gives rise to Foucault currents in the metal body of the satellite and produces moments retarding rotation of the satellite around the mass centre, which have to be taken into account in investigating rotation of the light unstabilised satellite. Evidently in a number of cases the variation in the law of rotation due to the magnetic field may have an effect, through variable orientation, on the resistance and consequently the life of the satellite.

Since the moments of inertia relative to the central axes of the rocket carrier of the first satellite differ greatly, then, on account of the fortuitous character of the initial disturbances, it can be assumed that the angle of nutation must be near to 90°. The motion represents a "somersaulting" condition. Observations confirm this conclusion. Analysis of experimental results from investigations of solar radiation afford conclusions regarding the motion of the second satellite around the mass centre at the beginning of its existence (1st November 1957). The parameters of motion around the mass centre are as follows: precession period 206 ± 10 sec; precession rate $1{\cdot}7°$/sec; angle of nutation $86 \pm 1°$; angle between the vector of kinetic moment and direction of the sun $90 \pm 5°$.

Optical observations were taken of the brightness of the second satellite at various points in January 1958. A sharp variation in intensity is readily explained by the fact that the condition of motion around the mass centre is near the somersault. The precession period of the satellite will be equal to twice the period of brightness variation of the satellite, and

optical observations of the second satellite agree with the period stated above. For the rocket carrier of the first satellite the precession period according to optical observations is approximately 80 sec. For the rocket carrier of the third satellite the precession period according to optical observations is less than 15 sec.

Orientation of the third satellite was determined by means of data obtained from treating the indications of the instruments fixing the variation in the position of the magnetometer frame during the passage of time. Preliminary results show that the conditions of motion of the third satellite were also near to those of the pure somersault. The precession period was equal to 125–135 sec. The period of self-rotation was considerably greater.

The orbital elements of the satellites were determined from results of treating a large number of observations; for the first satellite: 60,000 radio measurements and approximately 400 optical measurements; for the second satellite: 12,800 radio measurements and 2,000 optical measurements, and in addition 124 photographs were employed; for the third satellite on 7th July 1958 there were 52,750 radio measurements and 1,260 optical measurements.

Treatment of the data was carried out on high-speed computers. Table 1 shows the orbital parameters at the beginning of the path.

The orbits of satellites travelling in the upper atmosphere are suitable for observations with the exception of small zones near to the poles.

The curves of Fig. 2 show variations in the circling time of satellites around the earth. The curves of Fig. 3 show variations in minimum height (perigee), h_{min}, and Fig. 4 variations

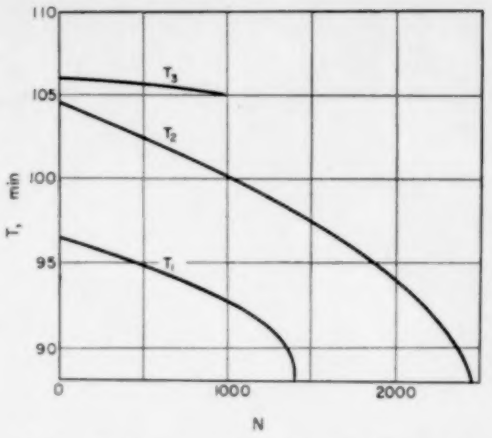


Fig. 2. Circling time of satellites as a function of the number of revolutions.

in maximum height of the trajectory h_{max} (apogee) for each of the three satellites.

The process of treating and predicting orbits is complicated and tedious, since it is necessary to work on a large number of measurements to obtain a high level of accuracy.

Since prediction is essential immediately after obtaining the measurement data, a special method of working out the observations is very important in order to provide rigorous ephemeris with sufficient speed. Methods of rigorous operational forecasting have been developed and successfully applied in practice.

Observation data regarding the evolution of the satellite orbit can be used for determining the density of the atmosphere. From the equations of motion of the centre of gravity of the satellite the following approximate formula can be derived for ΔT describing the variation

Table 1

Satellite	First	Second	Third
Circling time T (min)	96·17	103·75	105·95
Minimum height h_{min} (km)	226–228	225	226
Maximum height h_{max} (km)	947	1671	1881
Inclination of orbit i	65°, 129	65°, 310	65°, 188
Daily longitudinal drift of the ascending station $\Delta\Omega$	–3°, 157	–2°, 663	–2°, 528
Daily drift of the perigee distance from the angle $\Delta\omega$	–0°, 432	–0°, 407	–0°, 326
Daily drift of the period ΔT (sec/day)	1·8	3·08	0·75

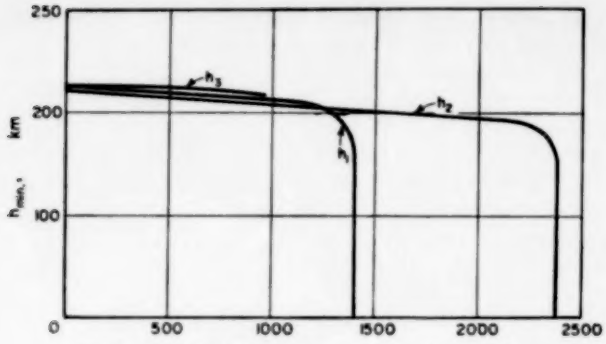


Fig. 3. Height of perigee as a function of the number of revolutions.

in the period of rotation of the satellite during one revolution:

$$\Delta T = -\frac{3\pi S a^2}{\sqrt{(\mu m)}} C_x \rho_\pi \sqrt{\left(\frac{2\pi H}{e}\right)} f\left(e, \frac{H}{ae}\right). \quad (3)$$

where $\mu = 3{\cdot}98600 \times 10^{20}$ cm^3/sec^2; m is the mass of the satellite; a is the large semi-axis of the osculating ellipse; $f(e, H/ae)$ is a known function of its arguments; e is the eccentricity determined at each instant in time from the data of the orbit; ρ_π is the density at the perigee.

For small values of eccentricity e and parameter $H/ae < \frac{1}{4}$ the asymptotic value of the function f is equal to unity[6].

The value ΔT, eccentricity e and semi-axis a are determined from observations, whilst the value H can be evaluated from the temperature of the atmosphere in the region of the perigee. Employing this, together with theoretical data regarding the coefficient C_x, it is possible to determine the density ρ_π. The accuracy of this calculation is related to the reliability of the values assumed for C_x and H.

Determination of the value H by means of orbit data is also possible, but for this purpose it is necessary to determine a very small value Δh_{min} with a high degree of accuracy, which renders a reliable result difficult.

For $H = 25$ km and $C_x = 0{\cdot}525$ the following densities are obtained for $h = 225$ km[8,9]:

$$\rho_\pi = 3 \times 10^{-13} - 4 \times 10^{-13} \text{ g/cm}^3$$

The densities obtained must be regarded as mean air-density values at the latitude of the perigee according to several regularly distributed points along the parallel.

In the calculations carried out employing data of the evolution of the orbit, attention is drawn to the systematic reduction in the fixed height of the product $C_x \rho_\pi \sqrt{H}$ with changeover of the perigee as a result of regression of the orbit from day to night. This effect is observed on all satellites. Hypotheses have been put forward that this is governed by a reduction in the atmospheric density at 225 km on changing over from day to night.

The density values set out above conform closely with density data obtained from rocket measurements.

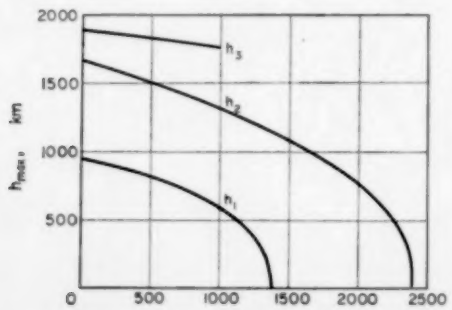


Fig. 4. Height of apogee as a function of the number of revolutions.

REFERENCES

1. D. E. Okhotsimskii, T. M. Eneyev and G. P. Taratynova, *U.F.N.* **63,** 33 (1957).
2. G. P. Taratynova, *U.F.N.* **63,** 51 (1957).
3. I. M. Yatsunskii, *U.F.N.* **63,** 59 (1957).

4. H. G. KRAUSE, *Secular and periodic disturbances in the path of an artificial earth satellite* in *Proc. 7th Intern. Astron. Congr., Rome 1956*, p 523.
5. J. DE SIKE, *The effect of the earth's oblateness and atmosphere on a satellite orbit* in *Proc. 7th Intern. Astron. Congr., Rome 1956*, p. 507.
6. R. JASTROW and C. A. PEARSE, *J. Geophys. Res.* **63**, No. 3 (1957).
7. V. V. BELITSKII, *Isk. Sput. Zem.* No. 1 (1958).
8. P. E. EL'YASBERG, *Isk. Sput. Zem.* No. 1 (1958).
9. M. L. LIDOV, *Isk. Sput. Zem.* No. 1 (1958).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 244-247. Printed in Great Britain.

PRELIMINARY REPORT ON GEOMAGNETIC MEASUREMENTS ON THE THIRD SOVIET ARTIFICIAL EARTH SATELLITE

S. SH. DOLGINOV, L. N. ZHUZGOV and N. V. PUSHKOV

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli* No. 2, p. 50 (1958)

Geomagnetic measurements on artificial earth satellites were envisaged in the scientific research programme of the Soviet and U.S.A. National Committees of the International Geophysical Year. The scientific importance of these measurements has been touched upon in a number of papers by American authors[1]. Papers by Soviet authors dealing with this problem include the work by Pushkov and Dolginov[2].

In accordance with the scientific research plan for the International Geophysical Year the first geomagnetic measurements were carried out on the third Soviet artificial earth satellite.

Simultaneously with the geomagnetic measurements on the third Soviet artificial earth satellite other geophysical measurements were carried out having a direct bearing on magnetic and electrical phenomena in cosmic space. The large amount of apparatus installed on the satellite introduced additional difficulties with regard to accurate measurements of magnetic field strength, associated with magnetic deviation, introduced by the apparatus on board. As a result of a study of the conditions of magnetic measurements on the satellite it was decided to carry out the measurements by means of a magnetometer with magnetically saturated pick-ups, turning automatically into the direction of the field vector.

Magnetometers of this type are ratio instruments and are surpassed in accuracy by proton precession magnetometers, but have the following advantages:

(1) Magnetometers of this type are less sensitive to disuniformity of the magnetic field and alternating magnetic electrical interference. Field disuniformity and alternating interference affect only the stability and accuracy of operation and do not impede the formation of the signal.

(2) It is possible by means of these magnetometers not only to measure the field strength, but also to determine the satellite orientation at different moments in time, as is necessary in the treatment of the data of other experiments.

(3) This type of magnetometer can be more readily adapted to the telemetering system employed.

Without dwelling on the design details of the magnetometer (a description of which will be given in another article), we will say only that the instrument is a second-harmonic magnetometer.

As distinct from instruments of this type employed in aeromagnetic surveying, it allows measurements at any magnetic latitude, for any satellite orientation, is fully automatic, has a low power demand and is relatively light. There is also a special device providing information regarding satellite orientation measurements in space together with the nature of rotation.

The measurement range of the instrument is 48,000 γ. The response rate of the servo system is 40–45 deg/sec. An idea of the magnetometer sensitivity can be obtained from Fig. 1.

Fig. 2 illustrates magnetometer stability over a period. These curves show instrument readings in the observatory at field variation ΔT during a test period on the instrument (Greenwich time), measured on observatory variometers. The mean zero drift is in the order of 2 γ/h. The magnetometer null-point was determined by comparison with the proton magnetometer.

Preliminary treatment of the material from the observations is proceeding at the present

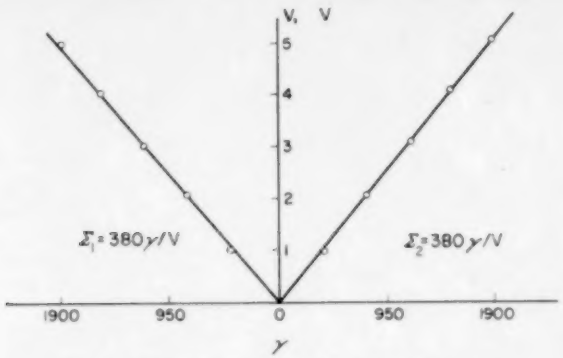


Fig. 1. Sensitivity Σ of the SG-45 magnetometer.

time. The geophysical interpretation of the results obtained will be presented after completing the treatment. The curves shown here illustrate the nature of the material obtained.

In Fig. 3, curves (1) and (2) show the readings of two potentiometer pick-ups, characterizing the angular displacements of the body of the satellite. From the curves showing the angular position of the satellite it is possible to obtain a numerical value of the relative variation in the position of the satellite in space for any moment in time. It follows from the curves that the satellite performs a precession movement within a period $T = 136$ sec. In addition, the satellite rotates around its own axis with a speed of approximately 0·36°/sec. There are grounds for assuming that in addition it is possible also to determine the absolute

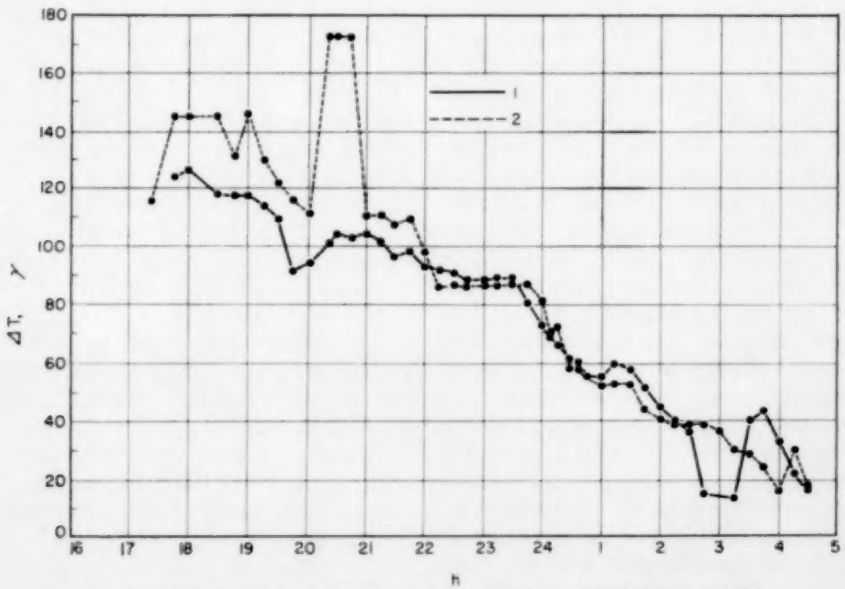


Fig. 2. Variation in intensity of the total vector Δ T, 11–12th March 1958.
1. Readings of SG-45 No. 29153.
2. Observatory magnetograph readings.

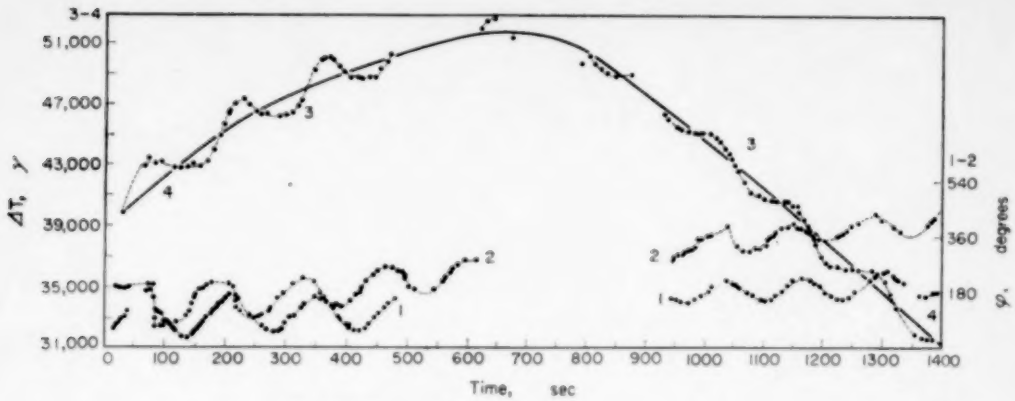


Fig. 3. Magnetometer charts.
Curves (1) and (2): variations in orientation.
Curves (3) and (4): magnetic field-strength variation.

orientation of the satellite in space relative to a given system of co-ordinates.

Curves (3) and (4) (Fig. 3) show magnetic field values directly measured by the magnetometer on a section of the trajectory extending over 23 min, in the altitude range 250–750 km. Curves (3) and (4) show that the magnetograms obtained on board the satellite are smoothly varying curves modulated by magnetic interference, closely associated with the body of the satellite. The period of modulation is equal to the precession period.

The influence of the level of magnetic properties of the apparatus on board was determined experimentally under laboratory conditions. The maximum magnetic deviation value is approximately 3000 γ. With such magnetic interference values it can be considered that the satellite magnetometer actually measures the magnetic field strength of the earth and the projection of the magnetic interference vector along the direction of the earth's field. As a result of precession of the satellite the value of the projection varies periodically both in magnitude and sign, with the result that the basic component of magnetic interference can be excluded.

Comparison of the measured magnetic field strength values relating to precession points of the rising and falling loops, provides a check on the null-point stability of the magnetometer over the period of operation of the apparatus. Good agreement (regarding the nature of the variation of the curves) is observed between the magnetograms of the satellite instrument and field strength curves constructed from the data of smoothed magnetic maps of ground values for points lying on the line of intersection of the orbital plane and the earth's surface. Comparison of these curves is employed for checking purposes in the initial treatment of the material.

A study of the preliminary material brings out certain points marked by relatively short period and rapid variations in the magnetic field. An example of such variations is shown in Fig. 4. They coincide in time with the passage of the satellite through the ionosphere zone. A particular point refers to the moment in time $10^h\ 46^m\ 11{\cdot}5^s$–$10^h\ 46^m\ 16^s$ 16th May 1958.

Similar phenomena are encountered on the photograms of material not yet treated. Their interpretation will be presented after studying all such cases.

A study of the experimental data obtained is proposed along the following lines:

(1) Comparison of field values measured by the magnetometer and calculated according to the potential theory. On the basis of calculations results will be set out of one of the last harmonic analyses of ground data.

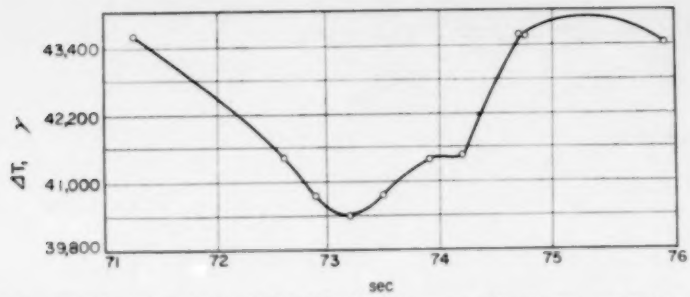


Fig. 4. Specific point obtained 16th May 1958, between $10^h\ 46^m\ 11^s$ and $10^h\ 46^m\ 16^s$.

(2) Comparison of the isolines of total magnetic field strength and cosmic ray intensity measured on the satellite.

(3) Analysis of the field above the eastern Siberian world magnetic anomaly, already studied considerably by ground, aeromagnetic survey and several times traversed by satellite, for the purpose of checking the hypotheses regarding the depth of its sources.

(4) Analysis of material for the purpose of clarifying the substance of the existence of flow systems in high atmospheric layers.

REFERENCES

1. *Scientific Uses of Earth Satellites*. University of Michigan Press, Ann Arbor (1956).
2. N. V. PUSHKOV and S. SH. DOLGINOV, *U.F.N.* **63**, 645 (1957).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 248-249. Printed in Great Britain.

THE OBSERVATION OF CORPUSCLES BY MEANS OF THE THIRD ARTIFICIAL EARTH SATELLITE

V. I. KRASOVSKII, YU. M. KUSHNER, G. A. BORDOVSKII, G. F. ZAKHAROV and E. M. SVETLITSKII

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli* No. 2, p. 59 (1958).

An observation of corpuscular streams was attempted on the third Soviet artificial earth satellite.[1] For this purpose use was made of indicators in the form of two fluorescent screens of ZnS(Ag) (2×10^{-3} g/cm^2), covered with aluminium foil 8×10^{-4} and 4×10^{-4} g/cm^2. Three aluminium diaphragms 5 mm thick were located before the screens with a diameter 5 cm, the intake orifice of which could trap corpuscles from a solid angle of $\frac{1}{2}$ steradians. The emission of the fluorescent screens was recorded by means of photo-electric multipliers. The electrical signals obtained were applied to the memory storage device and then transmitted by radiotelemetery to the earth.

The characteristic of the amplifying system was rendered non-linear in order to record the electrons falling perpendicularly over the whole fluorescent screen with an energy of 10^4 eV with a current $\sim 10^{-11} + 10^{-8}$ A/cm^2 in the case of a thick-foil indicator.

The thin foil had small micropores, allowing the passage of sunlight. Hence when directed towards the sun the indicator would give a reading in the middle of the scale when fitted with this foil.

With rotation of the satellite the light flux would have to be symmetrical relative to its maximum as the result of electromagnetic emission of the sun. A knowledge of these circumstances is necessary for evaluating the corpuscular action on a fluorescent screen with a thin foil. The thick-foil indicator did not give signal currents when directed towards the sun.

On 15th May 1958 intense signals were recorded from both indicators. Overshooting was observed at times and on infrequent occasions signals were absent around the threshold of sensitivity. Evidently the signal amplitude rises from the equatorial latitudes to higher latitudes and with altitude. Intense signals were several times observed with an abrupt start or abrupt cessation in approximately a second. The signal intensity varied throughout the entire time.

Treatment of the material obtained has not yet been completed, and it has not been compared with other observations. Final assessment is therefore not yet available.

An attempt may be made to explain the recorded signals either by irradiation of the fluorescent screens by ions (for example protons) or by X-ray emission, or by electrons, if all these agents possess an energy from a few kiloelectron volts to several hundred kiloelectron volts.

Since it was very difficult to assume a high power for the protons or X-ray emission, responsibility must be attributed to some agent associated with an extremely small energy flux. Hence hard electrons are not specifically the most striking.

Since the intensity of radiation of the fluorescent screen is somewhat greater with the thin foil than with the thick, it can be assumed that the electron energy does not exceed 10^4 eV and is possibly less than this value.

At the moment of overshooting, the electron flux exceeded 4×10^3 erg/sec cm^2 ster, and on the threshold of sensitivity was approximately 3 erg/sec cm^2 ster. A radiation intensity such as this doubtless complicates the investigation of X-ray emission of the sun and cosmic ν-radiation. The X-ray emission formed by the electrons may prove dangerous for living creatures travelling over a long period in the upper atmosphere. In addition, powerful electron fluxes may give rise to intense heating of the

upper atmosphere, raising its scale of altitudes. This is undoubtedly interesting in the light of much recent information regarding the upper atmosphere.

It is at present too early to announce any conclusive hypothesis regarding the origin of the corpuscles observed. We will limit ourselves to a few brief remarks. The usual delay time of geomagnetic disturbances relative to passage through the centre of the sun of any characteristic solar activity development does not support the assumption that electrons with an energy of a few kiloelectron volts up to several hundreds of kiloelectron volts could be primary corpuscles of the sun. It is just as difficult to assume that the electrons were formed near to the sun as a result of energy transformation of corpuscles-protons, travelling at speeds around 2×10^8 cm/sec, since an explanation of the energy fluxes quoted above (greater than 4×10^3 erg/sec cm^2 ster) would require a proton density of the primary fluxes of approximately 4×10^3, which has not so far been observed in studying the hydrogen emission of polar lights.

It is worth mentioning the possibility of explaining the observed phenomena on the basis of electron acceleration in the upper atmosphere (in the electrically conducting circuit formed along the lines of magnetic force at the contact with the earth's ionosphere) as the result of the varying magnetic fields congealed in the solar corpuscular stream. In this case it can be anticipated that the harder electrons will fall in the polar regions since a larger area contour is associated with the polar regions. The increase in electron fluxes during the day might be explained by the increase in ionisation at the boundary of the exosphere, as a result of which a greater quantity of ionised particles enters the upper atmosphere, or through the action of magnetic variations ρ_s, which are more intense during daytime. On achieving a certain velocity the electrons may perform oscillating movements along curved magnetic lines.

REFERENCES

1. V. I. Krasovskii Yu. M. Kushner and G. A. Bordovskii, *U.F.N.* **64**, 425 (1958).

Planet. Space Sci. Pergamon Press, 1961. Vol. 5, pp. 250-255. Printed in Great Britain.

ABSTRACTS

No. 4 (1960)

*M. D. KISLIK: **The path of an artificial earth satellite in the normal gravitational field of the earth** (pp. 3–17).

The problem is studied disregarding the influence of air resistance and gravity anomalies. The results can be employed in calculating the orbits of high-altitude artificial satellites and for the qualitative analysis of the influence of the earth's oblateness on the path of satellites. Formulae are developed providing a programme reducing the necessary computer time as against the programme based on numerical integration of the equations of motion. Relationships are provided for the height of an ellipsoidal artificial satellite above the surface of the common earth ellipsoid as a function of displacement of the satellite from the equator towards the pole. The theory provides closer approximation with regard to the difference due to the ellipsoidal character of the earth as compared with a point mass. In addition to purely calculation applications, the relationships obtained are applicable also to secular and periodic perturbations of the elements of the elliptical orbits of artifical satellites under the influence of the earth's oblateness. Investigations along these lines will be published in due course.

P. E. ELYASBERG and V. D. YASTREBOV: **Determination of the density of the upper atmosphere from the results of observations of the flight of the third Soviet artificial earth satellite** (pp. 18–30).

Observation of each individual artificial satellite yields extremely limited material regarding the state of the upper atmosphere, in accordance with the fairly narrow intervals of height variation and other conditions (illumination, solar activity, etc.), governing the air density at high altitudes. In view of the importance of accumulating data for the construction of a more complete dynamic model of the atmosphere, taking account of other factors than merely altitude, results of upper-atmosphere density determinations are presented as obtained from the more precise measurements of the orbit elements of the third Soviet earth satellite during the initial period of its existence. The orbit was determined considering the satellite as a material point, moving under the influence of the earth's gravity and air resistance. The elements of the satellite orbit were determined from measurement data as the result of the solution of the boundary-value problem for the system of differential equations of motion. Corrections to the initial conditions of motion were calculated by the method of least squares. Air-density values are tabulated for an altitude of 246 km. Discrepancy ranges are calculated analytically.

G. A. KOLEGOV: **Density variations in the upper atmosphere from data of fluctuations in the circling time of artificial earth satellites** (pp. 31–34).

Observations of artificial earth satellites have shown that the rate of diminution of the circling times, as the result of retardation in the upper atmosphere, does not increase in monotonous fashion as the satellite falls, but is subject to marked fluctuations. Since the rate of diminution is governed by the density of the upper atmosphere, an assessment of atmospheric density variations can be drawn from the speed fluctuations. Factors involved include the effect of solar heating, tidal effects in the atmosphere and variations in solar activity. The paper develops experimental results from observations of Soviet artificial earth satellites together with an interpretation of the results. Curves show density variations from earth-satellite observations, displaying a 12-hr (solar heating) trend alongside an 8-hr trend.

*I. M. YATSUNSKII: **Determination of illumination conditions and periods of an artificial satellite in shadow and in sunlight** (pp. 35–42).

For purposes of solar batteries, satellite visibility, thermal regime studies and orbit measurements, as well as for observations of the sun, a knowledge is necessary of the periods in shadow and in sunlight. A graphic-analytic method, taking account of the features of satellite movement in relation to the earth's shadow, has been found useful in projecting the early artificial satellites and in selecting the launching time. A brief analysis of the path of Soviet satellites Nos. 1, 2 and 3 in relation to the earth's shadow is illustrated by graphs showing visibility conditions of the third satellite from different latitudes, with periods in shadow and in the sun, and variations in height of the sun and dip of horizon for the second satellite in the early days of flight. The inclination of the plane of orbit of the Soviet satellites to the equator (65°) provides two periods of the year free of shadow. Beyond 800 km, brief shadow-free periods are possible throughout the year.

T. M. ENEYEV, A. K. PLATONOV and R. K. KAZAKOVA: **Determination of the orbit parameters of artificial satellites from ground-measurement data** (pp. 43–55).

The method, based on optical and radio observations and employing electronic computers, starts from a study of the cartesian co-ordinates x, y, z, stationary with reference to the stars, and with origin at the earth's centre. The x-axis is directed towards a point on the vernal equinox, the z-axis along the earth's axis of rotation and

the y-axis in such a way as to form a right-handed system of co-ordinates. The path of the satellite can be determined from a knowledge of x, y and z as a function of time and of the series of parameters characterizing the satellite orbit. Astronomical observations employ equatorial co-ordinates, and radio observations employ spherical co-ordinates. The orbit parameters are selected in such a way that the calculated co-ordinates are in agreement with the observed co-ordinates. For greater reliability the classic method of least squares was supplemented by a further method, providing convergence of the process of defining the orbit in the case of very rough first approximations. The length of forecast, i.e. the period for which sufficiently accurate ephemeris can be calculated, will depend both on the interval of time over which the observations are spread and on the quality of the observations.

G. P. TARATYNOVA: **Methods of numerical solution of equations of finite differences and their application to the calculation of orbits of artificial earth satellites** (pp. 56–81).

The methods are applicable to problems of celestial mechanics requiring the solution of systems of non-linear differential equations describing the path over long intervals of time, for example in determining the life of an artificial satellite, the evolution of the satellite path during its life and long-term forecasting of a satellite orbit. Use is made of the osculating elements of the orbit usually employed in celestial mechanics, and account is taken of the perturbing forces of atmospheric resistance, deviation of gravitational field of the earth due to oblateness and gravity anomalies. Two types of method are developed, of the Runge type and the Adams interpolation type. The Adams type methods require less calculations for a given increment of the sought function for each step, but demand a considerably greater memory storage when programming the algorithms on an electronic computer. Results are shown for orbit calculations on a satellite approaching the characteristics of the first Soviet artificial earth satellite based on 2-cycle solution methods. The integration step was at first 128 revolutions, diminishing with diminishing height. Calculations were carried out first for all the perturbing factors operating, then in the absence of atmospheric resistance and finally disregarding the earth's oblateness.

*A. I. LUR'E: **The equations for disturbed motion in the Kepler problem** (pp. 82–85).

The equations for disturbed motion of a planet were already known in part by Newton. The history of the problem and results of the equations were set out in Tisserand's *Celestial Mechanics* in 1889 and by Krylov in 1915 in his *Variation in the elements of elliptical orbits of Planets*. Tisserand, following the general methods of the theory of disturbed motion, calculated the Lagrange factors for the elliptical elements of the orbit. Krylov based his conclusions on geometrical constructions. The conclusion here proposed is based on the direct application of the method of variation of arbitrary constants to the expressions for the radius-vector of the planet and its speed vector, employing the known formulae for differentiation of unit vectors of the moving trihedral of axes. This result represents the most economical and simple of all those available.

K. P. STANYUKOVICH: **Elements of the theory of impact of solid bodies at high (cosmic) velocities** (pp. 86–117).

The problem of meteorite impact at cosmic velocities on the surface of a planet has a bearing on the study of micrometeorite collision at the surface of artificial earth satellites and space rockets. In the latter case the force of gravity can be disregarded in studying the scattering of the products of explosion from the crater. For simplification, the study is confined to the explosion of explosive charges of different powers and calorific values, for the purpose of examining the relationship of dimensions and shape of the crater and the nature of the discharge of explosion products to the characteristics of the medium and the gravitational acceleration on the given planet. The problem of meteorite explosion is thus reduced to the explosion of an equivalent quantity of explosive. The analysis is aimed at providing both cosmic and engineering conclusions, for the protection of satellites and rockets. A study of the ultimate effects of the ordinary explosion and the explosion of meteorites on solid surfaces yields the cosmic conclusion that small bodies, both of the solar system and in general, occurring in space, are continuously destroyed and disintegrate as the result of intercollisions.

B. A. MIRTOV: **Meteoric matter and some geophysical problems of the upper layers of the atmosphere** (pp. 118–134).

The scarcity of information and the difficulty of investigating finely-dispersed meteoric matter occurring above 100 km enhance the importance of investigations on artificial earth satellites carried out over a long period. Calculations provide particle size data at various altitudes from 70 to 130 km, considering dust of earthly origin. Data from the literature is tabulated for meteoric dust. The specific dust content for given altitudes is related to the rate of fall. The specific content of cosmic particles in the upper layers is different at different altitudes, suggesting that in the region 95 – 150 km the earth's atmosphere has some specific characteristics differing from other layers. Studying the problem from the standpoint of the dispersion of light on cosmic-dust particles, the diminishing number of particles with increasing particle size complicates the use of this method. Disuniformity of the luminosity of the night sky in 5577Å emission is attributed to the meteoric origin of this phenomenon. According to this hypothesis, the path of the disuniformities must lie predominantly from north-east to south-west. In the case of charged particles the latitude effect is explained by particles approaching the earth in large quanities at the magnetic pole. giving rise to intense luminescence in the polar regions. Material at present available favours the hypothesis of the meteoric

(micrometeoric) origin of luminosity of the night sky at 5577·3Å.

S. Sh. Dolginov, L. N. Zhuzgov and V. A. Selyutin: **The magnetometer apparatus of the third Soviet artificial earth satellite** (pp. 135–160).

The magnetometer was designed for automatic orientation regardless of the orientation of the satellite, for the purpose of magnetic-field-strength measurements. the four basic units are described. Saturated elements of the second-harmonic type were employed, made up of permalloy plates or wires carrying primary and secondary windings. Information was transmitted through five telemetering channels, two providing information as to the position of the orientation pickups, and the remaining three giving positive and negative magnetic-field variations, in addition to the position of the range switch. The sensitivity provided for an overall field variation on the two telemetering channels of $\pm 2{,}400\gamma$, with a range-switch in steps of $3{,}000\gamma$. Temperature curves and temperature-correction curves show only slight temperature variation and a temperature coefficient varying from 6γ/degree in the temperature range 0–20° to 2γ/degree in the range 20°–50°. The range switch of the measurement channel could accommodate a misalignment signal of $48{,}000\gamma$ in 1–1·5 sec. The consumption ranged from 12 W under stable conditions to 20 W with maximum performance. A specimen magnetometer record shows the gradual variation of the earth's magnetic field modulated by magnetic interference varying both in quantity and sign. Mean curves can be arrived at graphically or analytically. Diagrams illustrate the location of the instruments, the orientation system, the range-switch circuit and the magnetometer measurement circuit.

Ya. M. Shvarts: **A method for determining the electrical potential of bodies located in a plasma** (pp. 161–164).

The employment of rockets and artificial earth satellites for geophysical investigations necessitates the measurement of their electrical potential, since this affects the indications of the instruments. Several methods have been proposed and employed. The present paper introduces another approach to the problem of determining the potential of bodies located in a plasma, including rockets and satellites. This is based on the space-charge theory, which provides an expression for the electrostatic field strength at the surface of insulated bodies from the potential of the body, the thickness of the space-charge layer and two coefficients depending on the gas density. Limiting conditions are defined for bodies moving at high velocity, thus excluding the tail section where electrons mainly occur. Thus only data can be employed from electrostatic field-strength measurements and positive-ion current density measurements where the angle of inclination of the normal to the surface is not greater than 90° with respect to the course of the moving body.

*T. N. Nazarova: **Investigation of meteoric particles on the third Soviet artificial earth satellite.** (pp. 165–170).

Only integral characteristics and data obtained by extrapolation are available for meteoric particles with a mass of less than 10^{-4}g. Individual characteristics can be obtained only by direct methods by means of rockets and artificial satellites. Instruments employed include microphones, photomultipliers, fine-wire elements, accelerometer-type pickups and piezo-electric pickups. The third Soviet earth satellite employed an instrument for registering the number of particle impacts and numbers of pick-up impulses with explosion of meteoric particles at the surface. The instrument detected particles within the following ranges: $8\times10^{-9}-2{\cdot}65\times10^{-8}$g, $2{\cdot}65\times10^{-8}-1{\cdot}5\times10^{-7}$g, $1{\cdot}5\times10^{-7}-5{\cdot}6\times10^{-6}$g and greater than $5{\cdot}6\times10^{-6}$g. In the finest range, the signal was transmitted after the accumulation of thirty-two impacts, in the next two ranges after sixteen and four impacts, and in the last range after each impact. Assuming a mean particle speed of 40km/sec, impacts were recorded for particles with masses from 8×10^{-9} to $2{\cdot}65\times10^{-8}$g, possessing an energy in the order of $10^{4}-10^{5}$ ergs. A chart sets out the number of impacts recorded at altitudes of 1300–1500 km and 1700–1800 km. Variations in impact frequency were recorded during the period of the observations, the explanation of which requires further investigation.

V. G. Istomin: **Some results of positive-ion mass spectrum measurements on the third Soviet artificial earth satellite** (pp. 171–183).

During the period 15th–25th May some 15,000 mass spectra were obtained at altitudes from 225 to 980 km. Measurements were carried out in the northern hemisphere at latitudes 27°–65° between 7 and 11 a.m. Moscow time. Under these conditions ions of atomic oxygen O^+ predominate. Ions were also recorded with mass numbers 18: oxygen isotope $(O^{18})^+$, 28: molecular nitrogen N_2^+, 30: nitric oxide NO^+, and 32: molecular oxygen O_2^+. The composition of the ionosphere was found to vary as a function of altitude, with diminishing content of the heavy molecular ions NO^+, O_2^+ and N_2^+ with increasing altitude, together with increase in the light ion N^+. The molecular-oxygen content is less than 10^{-3} at altitudes above 500 km. The composition of the ionosphere is related to the latitude. There is a marked increase in atomic nitrogen ions at altitudes 225–350 km on passing from the 30–50° latitude to the 55–65° latitude. The relative content of molecular nitrogen, molecular oxygen and nitric-oxide ions is appreciably greater in the latitude range 55–65° than in the more southerly latitudes.

*Yu. G. Shafer and A. V. Yarygin: **Measurement of cosmic rays on geophysical rockets** (pp. 184–194).

In 1958 measurements of the global intensity of cosmic rays were carried out by means of identical Geiger counters and an ionisation chamber installed on a geophysical rocket, for the purpose of studying cosmic-ray intensities up to 200 km. Halogen counters were employed, with dimensions 18×150 mm and steel-wall thickness 0·1 mm. The instrument was aligned along the axis of the rocket. Tabulated data show the relationship of global cosmic-ray intensity to altitude, ranging from 6 km to 207 km, corresponding to a particle intensity of 0·21 particles/cm² sec to 1·40 particles/cm² sec. Secondary particles account for some 20 per cent of the total. A determination was made of the mean specific ionisation of primary particles in 1 cm³ argon at normal pressure, exceeding the mean ionisation of charged relativistic particles in the ratio of 1·8. The measured specific ionisation through a 1-cm path in argon at normal pressure was 189·1—8·2 cm⁻¹. The results also confirm the high level of reliability of the ionisation chamber and electronic units together with their insensitivity to supply-voltage variations within 20 per cent.

I. S. Shklovskii: **The artificial comet as a method of optical observation of space rockets** (pp. 195–204).

Astronomical observations of space rockets, even at lunar distances, is possible, if at all, only on a very few special telescopes. Methods for improving visibility include the balloon, which is extended pneumatically to 30 to 50m diameter. This, however does not provide anything like the necessary increase in visibility. A bright source of light momentarily switched on is also ruled out because of the high energies necessary. The natural phenomenon displayed by the comet is drawn on to provide a solution based on the resonant scattering of solar light by various molecules (C_2, OH, CN, etc.) and atoms (Na). Illumination of the gaseous tails of comets is also explained by scattering of the solar radiation by CO^+ and N_2^+ molecules. Calculation of the total quantity of molecules corresponding to illumination of the head and tail of a comet can be readily accomplished. A study of the efficacity of dust emanations yields the conclusion that the former method is less effective than sodium or lithium. The monochromatic radiation is also an advantage, permitting the use of light filters for enhancing clarity. The only advantage of the dust comet is its longer duration.

No. 5 (1960)

*L. I. Sedov: **Orbits of space rockets in the direction of the moon** (pp. 3–15).

The text of a survey presented to the annual meeting of the American Rocket Society in November 1959 provides information relating to the three Soviet space rockets. The most suitable angle of inclination of the plane of the first section of the orbit to the equatorial plane was determined at 65°. The most favourable conditions of observation from the U.S.S.R. require a flight time from the earth to the moon in the order of 0·5, 1·5, 2·5 or 3·5 days. Trajectory characteristics are set out for establishing the influence of initial parameters on the trajectory. The most suitable trajectory takes a path approaching the moon to a distance of 5,000–20,000 km, when the lunar attraction results in a trajectory which returns the object towards the earth to follow a satellite path with an initial perigee around 40,000 km and an apogee of 500,000 km, the return towards earth occurring in the higher latitudes of the northern hemisphere. Factual data of the three Soviet space rockets are set out. In the case of the third space rocket, favourable photographic conditions were achieved from the standpoint of illumination of the invisible face of the moon, two-thirds of which was photographed. Direct lighting of the surface was a disadvantage due to the absence of shadow from surface irregularities. Three specimen photographs are shown. Two of these are from originals, treated for radio-interference, whilst the third represents the first attempt at a composite photograph. The exercise is viewed from the standpoint of a new method in modern experimental astronomy.

S. Sh. Dolginov, E. G. Eroshenko, L. N. Zhuzgov, N. V. Pushkov and L. O. Tyurmina: **Magnetic measurements on the second space rocket** (pp. 16–23).

The measurements were aimed at examining the existence of a dipolar magnetic field on the moon, and collecting data concerning the upper corpuscular zone (radiation belt) of the earth, following on the earlier information obtained from the first space rocket. The problems are important in elucidating the earth's magnetism and the magnetism of cosmic bodies, in addition to clarifying magnetic disturbances and their relationship with solar activity, polar lights and cosmic rays. The dimensions of the upper zone, the radiation intensity and the distance of the zone of maximum radiation from the earth vary in relationship to magnetic storms and polar lights, indicating that the upper corpuscular zone is formed by solar corpuscles, detained by the magnetic field. Tabulated data illustrate the relationships. It seems also that the upper corpuscular zone forms a separate intermediate reservoir of solar corpuscles, which is replenished during magnetic disturbances and is gradually exhausted between disturbances. This may explain why polar lights are observed daily and not only when magnetic disturbances occur. Magnetometer readings near the moon point to the absence of a lunar magnetic field. Calculations show a maximum possible magnetic field strength at the moon's surface of one four-hundredth that at the earth's surface.

S. N. VERNOV, A. E. CHUDAKOV, P. V. VAKULOV, YU. I. LOGACHEV and A. G. NIKOLAEV: **Radiation measurements during the flight of the second space rocket** (pp. 24–29).

The apparatus was designed for obtaining fresh data of the upper radiation belt of the earth. for recording cosmic radiation along the path from the earth to the moon and for detecting a lunar radiation belt if such exists. Preliminary conclusions are drawn from measurements taken between 8000 and 120,000 km from the earth's centre and in the vicinity of the moon, starting at 40,000 km from the moon's surface. Regarding the composition of radiation in the upper radiation belt of the earth, two separate energy groups of particles are indicated, namely electrons with an energy of approximately 20 keV and electrons with an energy in the order of 2 MeV (or protons with an energy in the order of 10 MeV). The first group may be connected with thermodynamic equilibrium between protons and electrons in entering the earth's magnetic field. The second group would involve a disequilibrium process. In the vicinity of the moon, no increase in radiation intensity was observed up to a distance of 1000 km from the surface. Further observations between 1000 km and 0 km, though rendered more difficult due to the brief duration of this section of the flight, did not indicate any appreciable increase in intensity. Calculations point to a maximum possible magnetic field at the moon's surface one thousandth that at the earth's surface. Cosmic-ray measurements on all the instruments showed a constant intensity from 70,000 km from the earth's centre up to arrival in the vicinity of the moon.

L. V. KURNOSOVA, V. I. LOGACHEV, L. A. RAZORENOV and M. I. FRADKIN: **Investigation of cosmic rays during the flight of the second space rocket to the moon** (pp. 30–37).

One of the Cerenkov counters employed recorded the particles, and the other recorded nuclei with charges greater than 5 and 15. Both counters had cylindrical plexiglass detectors 2·6 cm diameter and 2·6 cm long. Charged particles would provide a splash of light if their speed exceeded 0·67 the speed of light, corresponding to a total energy of $1{\cdot}3\times10^9$ eV/C in the case of nuclei. In addition to the nucleus-count channels, one channel was provided for observation of the intensity of all the charged particles in the radiation belt, employing the principle that the photomultipliers of the Cerenkov counters are sensitive to X-rays penetrating the container shell under the influence of the charged particles. The mean flux of the different groups of nuclei does not vary at great distances from the earth, although deviations were recorded on individual sections of the path corresponding to short intervals of time. Whether these are statistical deviations or whether they represent variations in the nuclear component over a period of time is not explained. A rise in the α-particle counting rate in the range 10–30,000 km from the earth's centre is attributable either to a true increase in the α-particle flux (perhaps due to the accumulation of α-particles along orbits around the earth) or to the influence of intense radiation on the equipment in the radiation belt. Further investigations are necessary for clarifying the position.

T. N. NAZAROVA: **Results of an investigation of meteoric matter by means of instruments on space rockets** (pp. 38–40).

Instruments similar to those employed in the third Soviet artificial earth satellite were installed on the space rockets launched on the 2nd January, 12th September and 4th October, 1959. On the first space rocket the equipment was designed to record particles having masses in the ranges $2{\cdot}5\times10^{-9}$–$1{\cdot}5\times10^{-8}$g, $1{\cdot}5\times10^{-8}$–2×10^{-7}g and greater than 2×10^{-7}g. The signal was telemetered after 16 impacts in the most sensitive range, after 4 impacts in the next, and after each impact in the coarsest range. Each individual impact was recorded by the instruments installed on the second and third space rockets. Tabulated results for the three space rockets and for the third artificial earth satellite show that the density of meteoric matter is not constant in the earth's vicinity, varying both in time and space. During the flight of the second and third space rockets the greatest number of recorded discharge impulses referred not to the minimal impulses which could be recorded by the intrument, but to high-value impulses.

*B. V. RAUSHENBAKH and E. N. TOKAR: **Some control problems in interplanetary space** (pp. 41–53).

A study of space-missile rotation and course-holding starts from the assumption of n stabilising flywheels in the body of the missile, with arbitrary location and orientation of the axes, but invariable in time. Two systems of co-ordinates are introduced for setting up the equations of motion, taking for the common origin a point 0, being the centre of the masses of the system comprising the body of the missile and the reactive flywheels. The one system has its axes parallel to the axes of a given inertial system and the second system is fixed in relation to the body of the missile. Equations of motion are developed for the stabilised missile employing three reactive flywheels. Following a study of the rotating systems on the basis of the three angles separating the axes, an example is worked for the case of small deviations of the missile from a set position, from which it is derived that the fluctuations for the missile around the three axes $0X$, $0Y$ and $0Z$ are independent.

O. V. GURKO: **Determination of visibility conditions of space rockets** (pp. 54–59).

For the purpose of improving visibility, apparatus was installed for forming an artificial "comet" at a given moment in time, with a brilliance sufficient for observation. The method for selecting the appropriate time employed a projection of the object and of the sun on a map drawn on a cylindrical projection. Illumination lines were constructed on a transparent material, together with site-angle lines. Knowing the time and date of the

observation it is possible to determine the night-time position. A chart illustrates the problem of selecting the time for generating the "comet", from the date of observation, the launching time and the trajectory. Space-rocket trajectories are shown, corresponding to lunar inclinations in steps of 3°, proving quite adequate for any given launching date. The method has been used for analysing visibility conditions for comets on the first and second Soviet space rockets. In all cases the observations confirmed the accuracy of the method.

A. D. DANILOV: **The formation of NO^+ in the upper atmosphere** (pp. 60–65).

An important part is played by the charging of atomic-oxygen ions in the upper atmosphere with molecular nitrogen, yielding NO^+ ions, with fairly rapid recombination, dissociating to form nitrogen and oxygen. An attempt is made to obtain the quantity and altitude relationship of NO^+ ion concentration, comparing calculated results with observation data. Assuming other processes to be negligible, the two reactions must be in equilibrium. The ratio of the charge-capture coefficient to the dissociation coefficient is assumed to be 10^{-4}. A curve shows the NO^+/O^+ relationship as a function of altitude to 400 km. Satisfactory agreement between the theoretical curve and experimental data confirms the original hypothesis. Experimental day-time and night-time values confirm the anticipated result that the NO^+/O^+ ratio increases threefold at night at 300 km. NO^+ ions are distributed in the atmosphere in the form of a layer with a maximum around 200–250 km. The electron concentration varies only slightly in this range, since the O^+ and N_2 altitude-relationship plays the main part in the NO^+ distribution. N_2 falls off in the 100–200 km range, whilst O^+ increases, as a result of which the NO^+ concentration increases with altitude. Above 250 km O^+ stays practically constant, and therefore NO^+ diminishes in proportion to the diminution in N_2.

L. N. KUPEROV: **Signal tracking of the third Soviet artificial earth satellite on Cape Chelyuskin** (pp. 66–70).

Observations were carried out from the 16th May to the 6th June 1958, covering the 13th–296th orbit, on a frequency 20,005 kc/s. The apparatus was designed for field-strength measurements on the comparison principle. Projections of the satellite path on the earth's surface were plotted, and the maps thus obtained provided an assessment of the probability of observing signals on passing over a given point. An isoline map shows occasions when signals were observed as a percentage of occasions when the satellite passed over a given point. Radio signals were observed at distances over 5000 km only when the satellite was travelling towards the 65° parallel and a line connecting the satellite and Cape Chelyuskin was passing over the sea. For distances over the sea, the zone of reduced tracking probability shows the same horizontal distances as the zero isolines for overland distances, i.e. around the 4500 km circle. Up to 5,000 km, signal propagation is direct, without reflection from the ground. Between 5000 and 8500 km only one reflection occurs. Median field strengths were calculated from the maximum of 0·8 uV/m, 22 min after passing the equator, when the satellite was in the zone of direct visibility below the *F*2 layer. Maximum values outside the visibility zone were 0·4 and 0·3 uV/m respectively above and below the *F*2 layer.

I. M. YATSUNSKII and O. V. GURKO: **Variation in the albedo of the first artificial earth satellite under the influence of external factors** (pp. 71–73).

In addition to its importance in tracking satellites, the variation in albedo due to atmospheric effects, meteoric matter and other physical factors has a scientific interest. The polished aluminium surface of the first artificial satellite originally possessed a mirror surface, providing conclusions from the 200 observations taken during its flight. Tabulated data show site angles and phase angles with number of observations over two periods from the 9th October to 31st October and the 1st November to 7th December. During October the satellite retained its mirror-reflection properties, providing observations for phase angles up to 140°. Deterioration in the surface would reduce the brightness by 2·5 stellar magnitudes. Under these conditions, observation by the AT-1 astronomical telescopes would be practically impossible. In November, observations at phase angles over 120° were completely absent. Conclusions can be drawn not only regarding the diminished albedo of the satellite, but also regarding the changed character of reflection (diffuse instead of mirror). Confirmation is provided by the larger relative number of blank observations for clear-firmament conditions in the second period (20 per cent in November as against 10 per cent in October). Statistical treatment of the visual observations over the two months of flight thus points to a change in the reflecting properties of the satellite under the influence of external factors.

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 257-273. Printed in Great Britain.

HYDROGEN ATOMS AND IONS IN THE THERMOSPHERE AND EXOSPHERE

D. R. BATES and T. N. L. PATTERSON
Department of Applied Mathematics, Queen's University, Belfast

(*Received* 4 *March* 1961)

Abstract—The distribution of atomic hydrogen in the thermosphere and exosphere is computed taking into account the upward flow which balances the escape flux. Because of the upward flow the number-density gradient is much steeper than it would be in a static atmosphere. Attention is drawn to the fact that the ratio of the amount of hydrogen above the 100 or 110 km levels to the amount of hydrogen above the 200 or 300 km levels is a sensitive measure of the temperature of the exosphere. The evidence on the absolute abundance of atomic hydrogen is examined. It is concluded that the number density at the 120 km level is probably about 5×10^5/cm^3. The $Ly\alpha$ absorption line at this level is beyond the linear part of the curve of growth.

Consideration is also given to the steady-state distributions of O^+ and H^+ ions. In the lower part of the exosphere the number density of O^+ ions falls with increase in altitude (the associated scale height being twice that of the O atoms) and the number density of H^+ ions rises at the same rate (as was first pointed out by Dungey). The altitude at which the number densities of O^+ and H^+ ions become equal is calculated on various assumptions regarding the temperature and hydrogen content of the exosphere. It is found to be about 1200 km when the temperature is 1250°K and the hydrogen content corresponds to the number density cited near the end of the preceding paragraph. The gradient of the predicted electron-density distribution at several Earth radii is much less than that deduced from whistler studies.

The passage from charge transfer to diffusive equilibrium is discussed in an Appendix.

1. INTRODUCTION

In recent years much attention has been paid to the properties of the thermosphere and exosphere. This is partly because they are of interest in themselves and partly because they are of interest also in connexion with a variety of studies—with studies, for example, on the Van Allen radiation belt, on whistlers, on v.l.f. emission and on the transmission of radio waves through the ionosphere (cf. Hines[1]).

Amongst the basic problems is the calculation of how the hydrogen atoms and ions are distributed in altitude. Several important papers on this have been published. However, there are two major effects which have not been taken into account fully.

As was emphasized by Bates[2] the distribution of hydrogen atoms must be markedly influenced by the escape from the exosphere: thus the escape necessitates an upward flow through the thermosphere which in turn necessitates that the density of hydrogen atoms falls off more rapidly than would be the case in hydrostatic equilibrium. This has in general been ignored.

The difference between the gravitational potentials of the electrons and the ions gives rise to an electrostatic potential. Though Dungey[3] describes the position correctly in his stimulating original paper* on the subject it is commonly believed that the electrostatic potential halves the effective weight of each species of ion (provided the electron and ion temperatures are the same). The belief requires that the force on an ion due to its charge depends on its mass. It is therefore untenable.

In the present paper calculations are carried out on the distributions of hydrogen atoms and

* Mange[4] also describes the position correctly in a note which has appeared since this was written.

ions making allowance for the escape and for the true action of the electrostatic potential. Brief consideration is also given to deuterium.

2. MODEL ATMOSPHERES

For convenience the analytic model atmosphere introduced by Bates[2] was used. In this the temperature at altitude z is taken to be

$$T(z)=T(\infty)\{1-a\exp(-\tau\zeta)\} \tag{1}$$

where a and τ are constants given by

$$a=\{T(\infty)-T(z_0)\}/T(\infty) \tag{2}$$

and

$$\tau=\frac{1}{T(\infty)-T(z_0)}\left(\frac{dT}{dz}\right)_{z=z_0} \tag{3}$$

z_0 being some standard altitude; and where ζ is the geopotential distance from z_0 to z given by

$$\zeta=\int_{z_0}^{z}\frac{g(z)}{g(z_0)}dz \tag{4}$$

$g(z)$ being the local acceleration of gravity. Supposing that there is diffusive equilibrium above level z_0, the number-density of a constituent X at level z may be shown by integration of the hydrostatic equation to be

$$n(z/X)=n(z_0/X)\left\{\frac{1-a}{\exp(\tau\zeta)-a}\right\}^{1+\gamma}\exp(\tau\zeta) \tag{5}$$

with

$$\gamma=\frac{1-a}{\tau H(z_0/X)} \tag{6}$$

$H(z_0/X)$ being the scale height indicated.

Taking z_0 as 120 km it was assumed that

$$n(z_0/N_2)=4{\cdot}4\times10^{11}/\text{cm}^3,$$
$$n(z_0/O_2)=8{\cdot}8\times10^{10}/\text{cm}^3,$$
$$n(z_0/O)=3{\cdot}4\times10^{10}/\text{cm}^3,$$
$$T(z_0)=380^\circ\text{K},$$
$$\left(\frac{dT}{dz}\right)_{z=z_0}=20^\circ\text{K/km}$$

and

$T(\infty)=1000^\circ\text{K}$, 1250°K, 1500°K or 2000°K.

Several values of the temperature at the base of the exosphere were considered because this parameter controls the rate of escape of hydrogen atoms and, in consequence, is of central importance in the investigation.

Table 1 gives the computed temperature and density distributions. Detailed comparison with the information derived from satellites will not be made. As an indication of the position it is sufficient to recall that King-Hele and Walker[5] have found that the observations taken between October 1957 and March 1960 indicate the following mass-densities: $3{\cdot}9\times10^{-13}$ gm/cm^3 at 200 km; $3{\cdot}7\times10^{-14}$ gm/cm^3 at 300 km; $3{\cdot}1\times10^{-15}$ gm/cm^3 (day) and $1{\cdot}0\times10^{-15}$ gm/cm^3 (night) at 500 km. It should not be concluded from these that the lower of the values taken for the temperature at the base of the exosphere must be dismissed from further consideration: thus the adopted fractional abundance of atomic oxygen and the adopted temperature-gradient at the standard altitude may be underestimates; and in any case, the structure of the thermosphere is not always the same (cf. Nicolet[6,7]).

3. DISTRIBUTION OF HYDROGEN ATOMS

3.1. *Calculations*

If a minor atmospheric constituent Y is escaping from the exosphere the altitude-distribution of its number-density through the thermosphere is governed by the equation

$$\frac{dn(z/Y)}{dz}=-n(z/Y)\left[\frac{1}{H(z/Y)}+\{1+\alpha(Y)\}\frac{d\log T(z)}{dz}\right]-\frac{S(Y)}{D(z/Y)} \tag{7}$$

in which $S(Y)$ is the escape flux, $D(z/Y)$ is the diffusion coefficient and $\alpha(Y)$ is the thermal diffusion factor (cf. Chapman and Cowling[8]).

Denoting the altitude of the base of the exosphere (which was taken to be 500 km) by z_x the escape flux may be written

$$S(Y)=B(Y)n(z_x/Y) \tag{8}$$

where $B(Y)$ is a velocity which may be com-

Table 1. Temperatures and densities in model atmospheres

Model ($T(\infty)$ in °K)	Altitude z (km)	Temperature $T(z)$ (°K)	Number densities (/cm³) $n(z/O)$	$n(z/O_2)$	$n(z/N_2)$	Mean density $\rho(z)$ (gm/cm³)
1000	120	380	$3{\cdot}4\times10^{10}$	$8{\cdot}8\times10^{10}$	$4{\cdot}4\times10^{11}$	$2{\cdot}6\times10^{-11}$
	140	674	$9{\cdot}6\times10^{9}$	$1{\cdot}2\times10^{10}$	$7{\cdot}4\times10^{10}$	$4{\cdot}4\times10^{-12}$
	160	828	$4{\cdot}8\times10^{9}$	$3{\cdot}9\times10^{9}$	$2{\cdot}6\times10^{10}$	$1{\cdot}5\times10^{-12}$
	180	909	$2{\cdot}9\times10^{9}$	$1{\cdot}5\times10^{9}$	$1{\cdot}1\times10^{10}$	$7{\cdot}0\times10^{-13}$
	200	952	$1{\cdot}9\times10^{9}$	$6{\cdot}8\times10^{8}$	$5{\cdot}6\times10^{9}$	$3{\cdot}5\times10^{-13}$
	250	990	$7{\cdot}3\times10^{8}$	$1{\cdot}1\times10^{8}$	$1{\cdot}1\times10^{9}$	$7{\cdot}6\times10^{-14}$
	300	998	$3{\cdot}0\times10^{8}$	$1{\cdot}8\times10^{7}$	$2{\cdot}3\times10^{8}$	$2{\cdot}0\times10^{-14}$
	350	1000	$1{\cdot}3\times10^{8}$	$3{\cdot}2\times10^{6}$	$5{\cdot}1\times10^{7}$	$6{\cdot}0\times10^{-15}$
	400	1000	$5{\cdot}4\times10^{7}$	$5{\cdot}9\times10^{5}$	$1{\cdot}2\times10^{7}$	$2{\cdot}0\times10^{-15}$
	450	1000	$2{\cdot}4\times10^{7}$	$1{\cdot}1\times10^{5}$	$2{\cdot}7\times10^{6}$	$7{\cdot}6\times10^{-16}$
	500	1000	$1{\cdot}0\times10^{7}$	$2{\cdot}1\times10^{4}$	$6{\cdot}4\times10^{5}$	$3{\cdot}1\times10^{-16}$
1250	140	700	$9{\cdot}3\times10^{9}$	$1{\cdot}2\times10^{10}$	$7{\cdot}2\times10^{10}$	$4{\cdot}3\times10^{-12}$
	160	901	$4{\cdot}6\times10^{9}$	$3{\cdot}8\times10^{9}$	$2{\cdot}5\times10^{10}$	$1{\cdot}5\times10^{-12}$
	180	1028	$2{\cdot}8\times10^{9}$	$1{\cdot}6\times10^{9}$	$1{\cdot}2\times10^{10}$	$7{\cdot}0\times10^{-13}$
	200	1109	$1{\cdot}8\times10^{9}$	$7{\cdot}6\times10^{8}$	$6{\cdot}0\times10^{9}$	$3{\cdot}7\times10^{-13}$
	250	1204	$7{\cdot}9\times10^{8}$	$1{\cdot}5\times10^{8}$	$1{\cdot}5\times10^{9}$	$9{\cdot}7\times10^{-14}$
	300	1234	$3{\cdot}8\times10^{8}$	$3{\cdot}5\times10^{7}$	$4{\cdot}1\times10^{8}$	$3{\cdot}1\times10^{-14}$
	350	1245	$1{\cdot}9\times10^{8}$	$8{\cdot}8\times10^{6}$	$1{\cdot}2\times10^{8}$	$1{\cdot}1\times10^{-14}$
	400	1248	$9{\cdot}5\times10^{7}$	$2{\cdot}2\times10^{6}$	$3{\cdot}6\times10^{7}$	$4{\cdot}3\times10^{-15}$
	450	1249	$4{\cdot}8\times10^{7}$	$5{\cdot}9\times10^{5}$	$1{\cdot}1\times10^{7}$	$1{\cdot}8\times10^{-15}$
	500	1250	$2{\cdot}5\times10^{7}$	$1{\cdot}6\times10^{5}$	$3{\cdot}5\times10^{6}$	$8{\cdot}4\times10^{-16}$
1500	140	716	$9{\cdot}2\times10^{9}$	$1{\cdot}2\times10^{10}$	$7{\cdot}1\times10^{10}$	$4{\cdot}2\times10^{-12}$
	160	949	$4{\cdot}5\times10^{9}$	$3{\cdot}8\times10^{9}$	$2{\cdot}5\times10^{10}$	$1{\cdot}5\times10^{-12}$
	180	1113	$2{\cdot}7\times10^{9}$	$1{\cdot}6\times10^{9}$	$1{\cdot}2\times10^{10}$	$7{\cdot}0\times10^{-13}$
	200	1227	$1{\cdot}8\times10^{9}$	$7{\cdot}9\times10^{8}$	$6{\cdot}2\times10^{9}$	$3{\cdot}8\times10^{-13}$
	250	1385	$8{\cdot}1\times10^{8}$	$1{\cdot}8\times10^{8}$	$1{\cdot}7\times10^{9}$	$1{\cdot}1\times10^{-13}$
	300	1451	$4{\cdot}2\times10^{8}$	$5{\cdot}1\times10^{7}$	$5{\cdot}5\times10^{8}$	$3{\cdot}9\times10^{-14}$
	350	1479	$2{\cdot}3\times10^{8}$	$1{\cdot}5\times10^{7}$	$1{\cdot}9\times10^{8}$	$1{\cdot}6\times10^{-14}$
	400	1491	$1{\cdot}3\times10^{8}$	$4{\cdot}9\times10^{6}$	$7{\cdot}0\times10^{7}$	$6{\cdot}9\times10^{-15}$
	450	1496	$7{\cdot}3\times10^{7}$	$1{\cdot}6\times10^{6}$	$2{\cdot}6\times10^{7}$	$3{\cdot}2\times10^{-15}$
	500	1498	$4{\cdot}2\times10^{7}$	$5{\cdot}3\times10^{5}$	$1{\cdot}0\times10^{7}$	$1{\cdot}6\times10^{-15}$
2000	140	733	$9{\cdot}0\times10^{9}$	$1{\cdot}2\times10^{10}$	$7{\cdot}0\times10^{10}$	$4{\cdot}1\times10^{-12}$
	160	1008	$4{\cdot}3\times10^{9}$	$3{\cdot}8\times10^{9}$	$2{\cdot}5\times10^{10}$	$1{\cdot}5\times10^{-12}$
	180	1222	$2{\cdot}6\times10^{9}$	$1{\cdot}6\times10^{9}$	$1{\cdot}2\times10^{10}$	$6{\cdot}9\times10^{-13}$
	200	1389	$1{\cdot}7\times10^{9}$	$8{\cdot}3\times10^{8}$	$6{\cdot}3\times10^{9}$	$3{\cdot}8\times10^{-13}$
	250	1664	$8{\cdot}1\times10^{8}$	$2{\cdot}2\times10^{8}$	$1{\cdot}9\times10^{9}$	$1{\cdot}2\times10^{-13}$
	300	1814	$4{\cdot}5\times10^{8}$	$7{\cdot}3\times10^{7}$	$7{\cdot}3\times10^{8}$	$5{\cdot}0\times10^{-14}$
	350	1896	$2{\cdot}7\times10^{8}$	$2{\cdot}8\times10^{7}$	$3{\cdot}1\times10^{8}$	$2{\cdot}3\times10^{-14}$
	400	1941	$1{\cdot}7\times10^{8}$	$1{\cdot}1\times10^{7}$	$1{\cdot}4\times10^{8}$	$1{\cdot}2\times10^{-14}$
	450	1966	$1{\cdot}1\times10^{8}$	$4{\cdot}7\times10^{6}$	$6{\cdot}5\times10^{7}$	$6{\cdot}2\times10^{-15}$
	500	1981	$7{\cdot}1\times10^{7}$	$2{\cdot}0\times10^{6}$	$3{\cdot}1\times10^{7}$	$3{\cdot}5\times10^{-15}$

puted using equation (20) of the well-known review by Spitzer[9]. Some values are given in Table 2.

If its dependence on the constitution of the air is ignored the diffusion coefficient may be expressed approximately as

$$D(z/Y)=\frac{c(Y)\,T(z)^{\frac{1}{2}}}{n(z/\Sigma)} \qquad (9)$$

where $n(z/\Sigma)$ denotes the total number-density of all species and $c(Y)$ is a constant. From the

Table 2. Escape of hydrogen and deuterium

	Temperature $T(\infty)$ (°K)			
	1000	1250	1500	2000
B (H) (cm/sec)	$8{\cdot}2\times10^{2}$	$3{\cdot}1\times10^{3}$	$7{\cdot}4\times10^{3}$	$2{\cdot}2\times10^{4}$
B (D) (cm/sec)	1·0	$1{\cdot}5\times10^{1}$	$8{\cdot}8\times10^{1}$	$8{\cdot}2\times10^{2}$

atomic diameters involved it is estimated that $c(H)$ is some $2{\cdot}_1\times10^{15}$ cm/sec. The corresponding values of $c(D)$ is $1{\cdot}_2\times10^{15}$ cm/sec.* As far as can be judged the thermal diffusion factor α is about $-0{\cdot}25$ for both hydrogen and deuterium (cf. Grew and Ibbs[10]).

Using the model atmospheres described in section 2 and giving the various parameters the values indicated in the preceding paragraphs equation (7) was integrated downwards to the 120 km level starting from the 500 km level (the assumed location of the base of the exosphere). In view of (8) it is apparent that the value assigned to $n(z_x/Y)$ when carrying out the integration does not influence the form of the deduced distribution.

Mixing was assumed to keep the proportions of hydrogen and deuterium in the air constant between the 120 and 100 km levels. Above the 500 km level the distributions were obtained from the Opik–Singer ballistic theory (Opik and Singer[11] and Singer[12]).

The results of the calculations will now be presented. They have been scaled so that

$$n(500\ \text{km/H}) = 1\times10^{4}/\text{cm}^{3}.$$

The evidence on the absolute abundance of hydrogen will be discussed in the next Section.

Taking the independent variable to be

$$y = R_x/(R_x+z) \qquad (10)$$

where R_x is 6870 km, Table 3 gives the derived number densities, $n(y/\text{H})$, of hydrogen atoms in the exosphere. It also gives the corresponding mean scale heights, $\bar{H}_x(\text{H})$ defined according to

$$\bar{H}_x(\text{H}) = N_0(z_x/\text{H}) \Big/ n(z_x/\text{H}) \qquad (11)$$

* We wish to thank Dr. A. Dalgarno for a helpful discussion on the magnitudes of the diffusion coefficients.

Table 3. Distribution of hydrogen in the exosphere (Opik–Singer ballistic theory)

$y=R_x/(R_x+z)$	Temperature $T(\infty)$ (°K)			
	1000	1250	1500	2000
	Number density $n(y/\text{H})$ (per cm^3)			
1·0	$1{\cdot}0\times10^{4}$	$1{\cdot}0\times10^{4}$	$1{\cdot}0\times10^{4}$	$1{\cdot}0\times10^{4}$
0·9	$4{\cdot}8_5\times10^{3}$	$5{\cdot}4_5\times10^{3}$	$5{\cdot}8_3\times10^{3}$	$6{\cdot}1_6\times10^{3}$
0·8	$2{\cdot}3_4\times10^{3}$	$2{\cdot}9_7\times10^{3}$	$3{\cdot}4_3\times10^{3}$	$3{\cdot}9_6\times10^{3}$
0·7	$1{\cdot}1_0\times10^{3}$	$1{\cdot}5_8\times10^{3}$	$1{\cdot}9_7\times10^{3}$	$2{\cdot}4_7\times10^{3}$
0·6	$5{\cdot}0_3\times10^{2}$	$8{\cdot}0_9\times10^{2}$	$1{\cdot}0_9\times10^{3}$	$1{\cdot}4_8\times10^{3}$
0·5	$2{\cdot}1_8\times10^{2}$	$3{\cdot}9_1\times10^{2}$	$5{\cdot}6_4\times10^{2}$	$8{\cdot}3_6\times10^{2}$
0·4	$8{\cdot}6_3\times10^{1}$	$1{\cdot}7_3\times10^{2}$	$2{\cdot}6_8\times10^{2}$	$4{\cdot}3_0\times10^{2}$
0·3	$2{\cdot}9_6\times10^{1}$	$6{\cdot}5_9\times10^{1}$	$1{\cdot}1_0\times10^{2}$	$1{\cdot}9_2\times10^{2}$
0·2	7·7	$1{\cdot}9_2\times10^{1}$	$3{\cdot}4_3\times10^{1}$	$6{\cdot}5_5\times10^{1}$
$\bar{H}_x$ (H) (km)	1255	1600	1910	2340

Table 4. Distribution of hydrogen in the thermosphere

Altitude z (km)	Temperature $T(\infty)$ (°K)			
	1000	1250	1500	2000
	Number density $n(z/\mathrm{H})$ per cm^3			
100	$3{\cdot}6\times10^6$	$8{\cdot}4\times10^6$	$2{\cdot}6\times10^7$	$7{\cdot}4\times10^7$
110	$6{\cdot}8\times10^5$	$1{\cdot}6\times10^6$	$5{\cdot}0\times10^6$	$1{\cdot}4\times10^7$
120	$2{\cdot}0\times10^5$	$4{\cdot}7\times10^5$	$1{\cdot}5\times10^6$	$4{\cdot}2\times10^6$
130	$9{\cdot}3\times10^4$	$2{\cdot}2\times10^5$	$6{\cdot}4\times10^5$	$1{\cdot}8\times10^6$
140	$5{\cdot}7\times10^4$	$1{\cdot}3\times10^5$	$3{\cdot}5\times10^5$	$9{\cdot}6\times10^5$
150	$4{\cdot}1\times10^4$	$9{\cdot}0\times10^4$	$2{\cdot}2\times10^5$	$6{\cdot}0\times10^5$
160	$3{\cdot}2\times10^4$	$6{\cdot}5\times10^4$	$1{\cdot}5\times10^5$	$4{\cdot}1\times10^5$
180	$2{\cdot}3\times10^4$	$4{\cdot}1\times10^4$	$8{\cdot}8\times10^4$	$2{\cdot}3\times10^5$
200	$1{\cdot}8\times10^4$	$2{\cdot}9\times10^4$	$5{\cdot}6\times10^4$	$1{\cdot}4\times10^5$
250	$1{\cdot}4\times10^4$	$1{\cdot}8\times10^4$	$2{\cdot}8\times10^4$	$6{\cdot}0\times10^4$
300	$1{\cdot}3\times10^4$	$1{\cdot}4\times10^4$	$1{\cdot}8\times10^4$	$3{\cdot}3\times10^4$
350	$1{\cdot}2\times10^4$	$1{\cdot}2\times10^4$	$1{\cdot}4\times10^4$	$2{\cdot}1\times10^4$
400	$1{\cdot}1\times10^4$	$1{\cdot}1\times10^4$	$1{\cdot}2\times10^4$	$1{\cdot}5\times10^4$
450	$1{\cdot}0_5\times10^4$	$1{\cdot}0_5\times10^4$	$1{\cdot}1\times10^4$	$1{\cdot}2\times10^4$
500	$1{\cdot}0\times10^4$	$1{\cdot}0\times10^4$	$1{\cdot}0\times10^4$	$1{\cdot}0\times10^4$
	Fictitious number density $n'(z/\mathrm{H})$ per cm^3 in absence of upward flow (normalization as above)			
120	$3{\cdot}2\times10^4$	$3{\cdot}5\times10^4$	$3{\cdot}9\times10^4$	$4{\cdot}6\times10^4$

in which

$$N_0(z_x/\mathrm{H})=\int_{z_x}^{\infty} n(z/\mathrm{H})\,dz=R_x\int_0^1 n(y/\mathrm{H})\,y^{-2}dy. \tag{12}$$

The situation in the thermosphere is described by Table 4. As may be seen $n(z/\mathrm{H})$ is predicted to be a rapidly decreasing function of z in the lower part of the region. Up to the 120 km level this is an immediate consequence of the assumption that the air is mixed. However, the prediction does not depend on the validity of the assumption. It is apparent from the trend just above the 120 km level that a steep number-density gradient is in any event needed to keep the upward flow of hydrogen atoms equal to the escape flux. The extent of the effect is shown by the last row of Table 4 which gives

$$n'(120\,\mathrm{km}/\mathrm{H})=n(500\,\mathrm{km}/\mathrm{H})\left\{\frac{T(500\,\mathrm{km})}{T(120\,\mathrm{km})}\right\}^{(15/16)+s}\times\left\{\frac{n(120\,\mathrm{km}/\mathrm{O})}{n(500\,\mathrm{km}/\mathrm{O})}\right\}^{1/16}, \tag{13}$$

which is what the number-density of hydrogen atoms at 120 km would be if there were no upward flow (cf. Bates and McDowell[13]). As far as can be judged the effect must be real. If the steepness of the gradient is to be avoided it seems necessary to suppose that the value adopted for the constant of proportionality appearing in expression (9) for the diffusion coefficient is much too low; that the air densities have been seriously overestimated; or that the upward flow is greatly reduced by the entry of hydrogen atoms into the atmosphere from interplanetary space. These suppositions have little plausibility.

With the aid of Tables 3 and 4, $N_0(z/\mathrm{H})$, the number of hydrogen atoms in a vertical column of unit cross-sectional area with its base at altitude z may readily be computed for the various model atmospheres (Table 5). Almost all the hydrogen above the 100 km level is found to be quite close to this level or to be in the exosphere.

The time constants,

$$\tau(z)=N_0(z/\mathrm{H})/S(\mathrm{H}) \tag{14}$$

Table 5. Depth of atmospheric hydrogen*

Altitude of base of region z (km)	Temperature $T(\infty)$ (°K)			
	1000	1250	1500	2000
	$N_0(z/\mathrm{H}) = \int_z^\infty n(z/\mathrm{H})\,dz$ per cm²			
100	$4{\cdot}2\times10^{12}$	$8{\cdot}2\times10^{12}$	$2{\cdot}1\times10^{13}$	$5{\cdot}5\times10^{13}$
105	$2{\cdot}9\times10^{12}$	$5{\cdot}0\times10^{12}$	$1{\cdot}1\times10^{13}$	$2{\cdot}9\times10^{13}$
110	$2{\cdot}4\times10^{12}$	$3{\cdot}7\times10^{12}$	$7{\cdot}5\times10^{12}$	$1{\cdot}8\times10^{13}$
120	$2{\cdot}0\times10^{12}$	$2{\cdot}8\times10^{12}$	$4{\cdot}7\times10^{12}$	$9{\cdot}8\times10^{12}$
130	$1{\cdot}9\times10^{12}$	$2{\cdot}5\times10^{12}$	$3{\cdot}7\times10^{12}$	$6{\cdot}9\times10^{12}$
140	$1{\cdot}8\times10^{12}$	$2{\cdot}4\times10^{12}$	$3{\cdot}2\times10^{12}$	$5{\cdot}6\times10^{12}$
160	$1{\cdot}7\times10^{12}$	$2{\cdot}2\times10^{12}$	$2{\cdot}8\times10^{12}$	$4{\cdot}3\times10^{12}$
200	$1{\cdot}6\times10^{12}$	$2{\cdot}0\times10^{12}$	$2{\cdot}4\times10^{12}$	$3{\cdot}4\times10^{12}$
250	$1{\cdot}5\times10^{12}$	$1{\cdot}9\times10^{12}$	$2{\cdot}2\times10^{12}$	$2{\cdot}9\times10^{12}$
300	$1{\cdot}5\times10^{12}$	$1{\cdot}8\times10^{12}$	$2{\cdot}1\times10^{12}$	$2{\cdot}7\times10^{12}$
400	$1{\cdot}4\times10^{12}$	$1{\cdot}7\times10^{12}$	$2{\cdot}0\times10^{12}$	$2{\cdot}5\times10^{12}$
500	$1{\cdot}3\times10^{12}$	$1{\cdot}6\times10^{12}$	$1{\cdot}9\times10^{12}$	$2{\cdot}3\times10^{12}$

*The normalization is as in Table 4.

Table 6. Time constants $\tau(z)$ defined by equation (14) of text

Altitude z (km)	Temperature $T(\infty)$ (°K)			
	1000	1250	1500	2000
	$\tau(z)$ (hr)			
100	140	75	80	70
110	80	35	28	22
120	70	25	17	12
150	60	20	11	6
200	55	18	9	4
300	50	16	8	3

which are of interest in connexion with the response of the hydrogen in the thermosphere to a change of temperature*, may be obtained by combining Tables 2 and 5. Some values are given in Table 6. Unless the temperature is low the response in the upper part of the thermosphere is quick enough to allow diurnal variations to be followed to a significant extent. It is probable too, that the brief intervals of high temperature postulated by Spitzer in connexion with the escape of helium could be followed closely.

Provided the altitudes z_1 and z_2 are suitably chosen the ratio

$$r(z_1/z_2) = N_0(z_1/\mathrm{H})/N_0(z_2/\mathrm{H}) \qquad (15)$$

is very sensitive to $T(\infty)$, as may be seen from Table 7 (which was prepared from the information contained in Table 5). A series of

* An increase in $T(\infty)$ leads to a decrease in the amount of hydrogen above the level at which mixing ceases to be important. (The constancy of $n(500\mathrm{km}/H)$ across Table 3 is, of course, merely due to the normalization procedure adopted.)

Table 7. Ratios $r(z_1/z_2)$ defined by equation (15) of text

Altitudes (km)		Temperature $T(\infty)$ (°K)			
		1000	1250	1500	2000
z_1	z_2	$r(z_1/z_2)$			
100	300	2·8	4·5	9·6	20·5
100	200	2·6	4·1	8·5	16·4
110	300	1·6	2·0	3·5	6·6
110	200	1·5	1·8	3·1	5·3

measurements of the ratio by rocket-borne instruments would be very instructive.

Deuterium escapes much less rapidly than hydrogen (cf. Table 2). Its abundance relative to hydrogen is therefore greater in the thermosphere than in the main atmosphere. According to the calculations which have been described the factor by which the relative abundance is increased at the 300 km level is 5 if $T(\infty)=1000°K$, 9 if $T(\infty)=1250°K$, 15 if $T(\infty)=1500°K$, and 8 if $T(\infty)=2000°K$.* These values depend on the assumption that 120 km is the altitude of the level up to which the air is kept mixed. They are therefore rather uncertain.

3.2. *Absolute abundance*

The free hydrogen in the atmosphere originates from water vapour and methane which rise from the Earth's surface and are dissociated by photo-chemical action (Bates and Nicolet[14]). A study of the source does not provide a reliable estimate of the amount present.

Kupperian *et al.*[15] have found that the Earth has an albedo of 0·42 towards the *Ly*α radiation, λ1215·7 from the night sky and an albedo of less than 0·02 towards the *Ly*α radiation direct from the Sun. Taking into account the fact that in addition to being scattered by atomic hydrogen the *Ly*α radiation is absorbed by the molecular oxygen Brandt and Chamberlain[16] have calculated the night and day albedos to be expected in various circumstances. They expressed these albedos as functions of

$$\eta=2\cdot9\times10^{-18}\,\overline{n(O_2)/n(H)} \qquad (16)$$

in which the numerical factor is the ratio of certain absorption cross-sections and in which $\overline{n(O_2)/n(H)}$ is the mean ratio of the number-density of molecular oxygen to the number-density of atomic hydrogen in the lower part of the atmosphere.

It has been suggested that the night sky *Ly*α radiation is solar *Ly*α radiation scattered by interplanetary hydrogen (Kupperian *et al.*[15], Struve[17], Shklovsky[18]). In this hypothesis the orbital motion of the Earth would give rise to a Doppler shift. The calculations of Brandt and Chamberlain[16] show that owing to the shift η would have to be as small as 1×10^{-5} if the observations of Kupperian *et al.* are to be explained without invoking the possibility of an excitation process in the atmosphere. The corresponding abundance of hydrogen is unacceptably high: for example, if $n(120\,km/O_2)$ is $8\cdot8\times10^{10}/cm^3$ as in Table 1 the corresponding $n(120\,km/H)$ is about $3\times10^8/cm^3$.

Johnson and Fish[19] have argued that the scattering is by hydrogen in the exosphere. Provisionally assuming that this is correct and adopting a value of 6 for the ratio of the temperature in the exosphere to the temperature in the lower part of the atmosphere it is found from the work of Brandt and Chamberlain that η is about 2×10^{-2}; and hence, again taking $n(120\,km/O_2)$ to be as in Table 1, that $n(120\,km/H)$ is some $1\times10^5/cm^3$. The relia-

* The last case differs from the other three in that in it alone the upward flow of deuterium is important.

bility of the estimate is difficult to assess. Corrections may be needed because

$$n(120\text{ km}/O_2)\Big/n(120\text{ km}/\text{H})$$

may differ from $\overline{n(O_2)/n(H)}$ and because the adopted value of $n(120\text{ km}/O_2)$ may be in error; but these corrections are likely to be opposite in sense. The main uncertainty arises from the possibility that the *Ly*α radiation from the night sky is not entirely due to the hydrogen in the exosphere but is partly due to the hydrogen in interplanetary space. Because of this possibility the value found for $n(120\text{ km}/\text{H})$ is strictly only a lower limit.

The upper bound which the measurements have set to the day albedo is not very informative. Taking cognizance of the absorption core (cf. Purcell and Tousey[20]) it is judged that all that can be deduced with assurance is that η is greater than about 2×10^{-5} and hence that $n(120\text{ km}/\text{H})$ is less than some $1\times10^{8}/\text{cm}^3$.

Using the results of Kupperian *et al.*[15] on the intensity distribution of the *Ly*α radiation from the night sky Johnson and Fish[19] have estimated that $N_0(120\text{ km}/\text{H})$ is about $7\times10^{12}/\text{cm}^2$. The estimate depends on simplifying assumptions of doubtful validity. A firm conclusion regarding the amount of hydrogen cannot be drawn until the relevant radiative-transfer problem is solved.

Direct evidence on the abundance has been reported by Purcell and Tousey[20] who have succeeded in obtaining micro-photometer tracings of the solar *Ly*α radiation at altitudes between about 100 and 200 km. The rocket concerned was launched from White Sands at 7 hr local time on 21 July 1959.

When interpreting the data it is necessary to take account of the van Rhijn effect. If ξ is the Sun's zenith angle and R_0 is the radius of the Earth the number of atmospheric hydrogen atoms per unit area long the line of view from altitude z_1 is

$$N_\xi(z_1/\text{H}) = \int_{z_1}^{\infty} \frac{n(z/\text{H})}{\{1-[(R_0+z_1)/(R_0+z)]^2\sin^2\xi\}^{1/2}}\,dz. \qquad (17)$$

Writing

$$n(z/\text{H}) = n(z_1/\text{H})\exp\{-(z-z_1)/\bar{H}(\text{H})\} \qquad (18)$$

where $\bar{H}(\text{H})$ is a local mean scale height and noting that

$$\bar{H}(\text{H}) \ll R_0 \qquad (19)$$

it may be shown that

$$N_\xi(z_1/\text{H}) = sN_0(z_1/\text{H}) \qquad (20)$$

in which

$$s = \Psi\left(\frac{2\bar{H}(\text{H})\tan^2\xi}{R_0+z_1}\right)\sec\xi \qquad (21)$$

with

$$\Psi(x) = (\pi/x)^{1/2}\exp(1/x)\,\text{erfc}\,(1/x^{1/2}) \qquad (22)$$

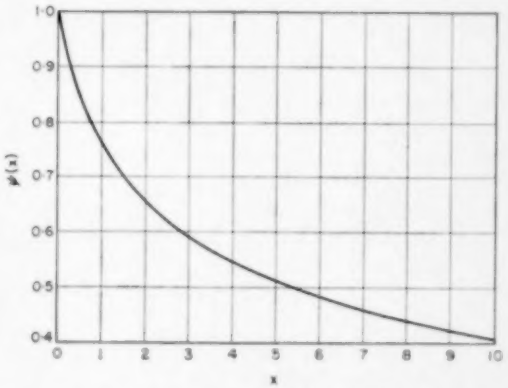


Fig. 1. Function $\psi(x)$ defined in equation (22) of text.

When the exposures of Purcell and Tousey were being made ξ was 69°. In the case of the hydrogen at the base of the thermosphere the corresponding value of s is 2·8; and in the case of the hydrogen in the exosphere it is 1·7 or 1·6 according as $T(\infty)$ is about 1000°K or is about 1250°K—the higher values of $T(\infty)$ need not be considered here since the flight took place in the early morning.

The relation between $N_\xi(z_1/\text{H})$, the number of hydrogen atoms per unit area along the line of view, and $W_\xi(z_1)$, the equivalent width of the absorption line cannot be expressed analytically except on the linear part of the curve of growth where the standard formula reduces to

$$N_\xi(z_1/\text{H}) = 1{\cdot}83\times10^{14}W_\xi(z_1) \qquad (23)$$

$W_\xi(z_1)$ being in angstroms. In general, the most convenient procedure is to treat the hydrogen as forming two layers, a hot (temperature T^u) upper layer and a cool (temperature T^L) lower layer. Let the contribution to $N_\xi(z_1/\mathrm{H})$ from these layers be N_ξ^u and N_ξ^L respectively; and let the equivalent width due to the upper layer alone be W_ξ^u (Fig. 2) and that due to the lower layer alone be W_ξ^L (Fig. 3). It is found empirically that the equivalent width due to the two layers together is approximately

$$W_\xi^{uL} = W_\xi^u + \beta(N_\xi^u) W_\xi^L \qquad (24)$$

with

$$\beta(N_\xi^u) = \exp(-1{\cdot}5 \times 10^{-13} N_\xi^u). \qquad (25)$$

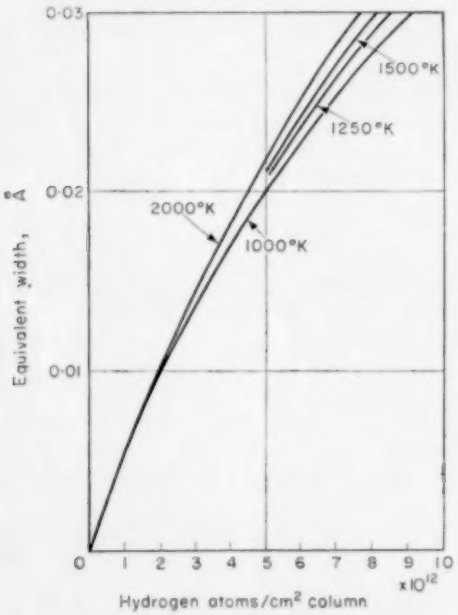


Fig. 2. Curves of growth for λ 1215·7. The temperature of the absorbing hydrogen is as indicated on each curve.

The equivalent width in any particular case of interest may now be readily obtained. Fig. 4 contains some results for the computed (Tables 4 and 5) distributions of hydrogen atoms. The measurements of Purcell and Tousey[29] gave $W_{69°}$ (100 km) to be about 0·026 Å. It would hence appear that if on the occasion of the flight $T(\infty)$ was 1000°K then $n(500\,\mathrm{km/H})$ was $1{\cdot}1_5 \times 10^4/\mathrm{cm}^3$, and if $T(\infty)$ was 1250°K

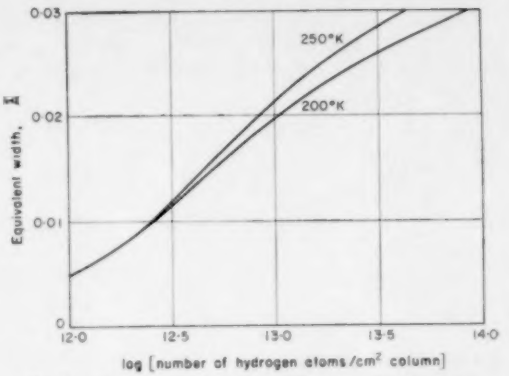


Fig. 3. Further curves of growth for λ 1215·7. The temperature of the absorbing hydrogen is as indicated on each curve.

then $n(500\,\mathrm{km/H})$ was $6{\cdot}_5 \times 10^3/\mathrm{cm}^3$. The deduced abundances are quite sensitive to the measured equivalent width: thus the values of $n(500\,\mathrm{km/H})$ corresponding to those just cited would be $1{\cdot}5_5 + 10^4/\mathrm{cm}^3$ and $9{\cdot}_5 + 10^3/\mathrm{cm}^3$ respectively if $W_{69°}$ (100 km) were 0·029 Å (which is about 10 per cent higher than the reported width).

We shall tentatively suppose that $n(100\,\mathrm{km/H})$ is about $8{\cdot}4 \times 10^6/\mathrm{cm}^3$ which corresponds to

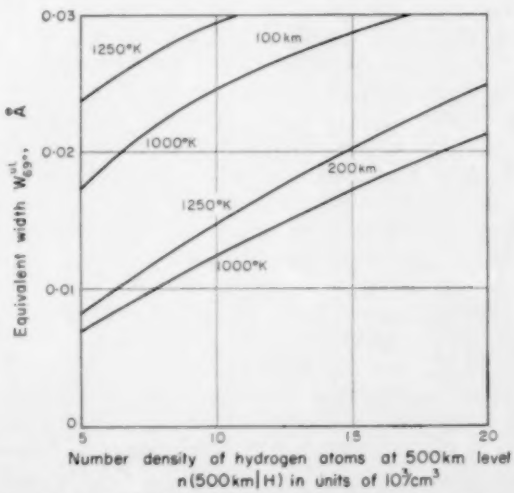


Fig. 4. Equivalent width of λ 1215·7 absorption line viewing at zenith angle of 69° from 100 km and 200 km levels. The distribution of hydrogen atoms is taken to be as in Tables 3 and 4 (with $T(\infty)$ either 1000°K or 1250°K).

$n(500\,\mathrm{km/H})$ being $2{\cdot}3\times10^4/\mathrm{cm}^3$, $1{\cdot}0\times10^4/\mathrm{cm}^3$, $3{\cdot}2\times10^3/\mathrm{cm}^3$ or $1{\cdot}1\times10^3/\mathrm{cm}^3$ according as $T(\infty)$ is 1000°K, 1250°K, 1500°K or 2000°K.

4. CALCULATION OF STEADY-STATE DISTRIBUTION OF IONS

4.1. *General remarks*

As was pointed out by Dungey[3] the main source and sink of hydrogen ions in the thermosphere is the charge transfer process

$$\mathrm{H}+\mathrm{O}^+ \rightleftharpoons \mathrm{H}^+ + \mathrm{O}. \qquad (26)$$

Noting that the energy balance is essentially exact it is seen that in equilibrium

$$n(z/\mathrm{H})\,n(z/\mathrm{O}^+)=\sigma n(z/\mathrm{H}^+)\,n(z/\mathrm{O}) \qquad (27)$$

where σ is the value, 8/9, of the ratio of the product of the statistical weights of the systems on the left of (26) to the product of the statistical weights of the systems on the right.

Process (26) is of the accidental or assymetrical resonance type. The rate coefficient for this type of process is not as high as for a process of the symmetrical resonance type like

$$\mathrm{H}+\mathrm{H}^+ \rightleftharpoons \mathrm{H}^+ + \mathrm{H} \qquad (28)$$

(Bates and Lynn[21]) but is unlikely to be very low. A reliable estimate of the rate coefficient of (26) is not available. Uncertainty might therefore be expected in the determination of the location of the boundary between the region in which there is charge transfer equilibrium (27) and the region in which there is diffusive equilibrium; and consequently, uncertainty might also be expected in the calculation of how the ions are distributed with altitude through the upper thermosphere and exosphere. However, as will be demonstrated, the distributions of ions corresponding to charge transfer equilibrium differ from the distributions corresponding to complete diffusive separation only in so far as the distributions of ions are affected by the rotation of the Earth and the distribution of hydrogen atoms is affected by escape. Consequently, uncertainty only arises in the steady state because of the possibility that there may be an intermediate zone where the number densities are neither high enough for charge transfer equilibrium nor low enough for complete diffusive equilibrium. This possibility is discussed in the Appendix. It is judged to be rather unlikely and will be neglected throughout the remainder of the paper. The neglect cannot cause serious error. Should the intermediate zone exist its effect would be to make the value of $n(z/\mathrm{H}^+)\Big/n(z/\mathrm{H})$ at any level in the diffusive equilibrium region somewhat less than the value corresponding to local charge transfer equilibrium.

Diurnal and other temporal changes occur in the F-layer. Since the response of the distant parts of the exosphere to these is clearly slow a true steady state may never be reached. It may be noted that process (26) keeps the change in $n(500\,\mathrm{km/H}^+)$ from being as great as might at first be expected: thus an increase in $n(500\,\mathrm{km/O}^+)$ is likely to be accompanied by a rise in $T(\infty)$ and hence by an increase in $n(500\,\mathrm{km/O})$ and by a decrease in $n(500\,\mathrm{km/H})$.

4.2. *Detailed formulation*

Suppose that the upper part of the F-layer contains electrons, e, of mass m_e and ions, O^+ and H^+, of mass m_{O} and m_{H} respectively*. Let z_s be the altitude of a reference level in the region where the ions are in diffusive equilibrium and measure potential energies with respect to this reference level at the equator. Assume that the region is isothermal and take $T(\infty)$ to be the temperature of the electrons, ions and neutral particles.

The effective gravitational potential energy of unit mass rotating with the Earth is given by

$$V_{g\Omega}(z)=V_g(z)+V_\Omega(z) \qquad (29)$$

with

$$V_g(z)=GM\left\{\frac{z-z_s}{(R_0+z_s)(R_0+z)}\right\} \qquad (30)$$

$$V_\Omega(z)=-\tfrac{1}{2}\Omega^2\{(R_0+z)^2\cos^2\theta-(R_0+z_s)^2\} \qquad (31)$$

in which G is the gravitational constant; M, R_0 and Ω are the mass, radius and angular velocity

* Satellite measurements (Istomin[22]) show that N^+ ions are sufficiently rare to be neglected in the present connexion.

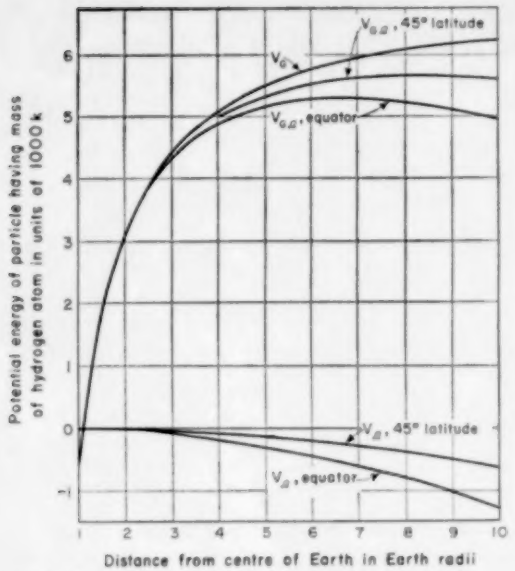

Potential energy of particle having mass of hydrogen atom in units of 1000 k
Distance from centre of Earth in Earth radii
$V_{G\Omega}$, 45° latitude
V_G
$V_{G\Omega}$, equator
V_Ω, 45° latitude
V_Ω, equator

Fig. 5. Gravitational and centrifugal potentials.

of the Earth; and θ is the latitude. Fig. 5 shows $V_G(z)$, $V_\Omega(z)$ and $V_{G\Omega}(z)$ plotted against

$$r \equiv (R_0 + z) \Big/ R_0 \tag{32}$$

for points on the equatorial plane and at 45° latitude.

The space-charge throughout the region is almost zero (Dungey[3]). Hence V_E, the electrostatic potential energy of unit positive charge must be such that in the steady state

$$n(z_s/e) \exp \frac{-1}{kT(\infty)} \{m_s V_{G\Omega} - eV_E\}$$
$$= n(z_s/\mathrm{O}^+) \exp \frac{-1}{kT(\infty)} \{m_\mathrm{O} V_{G\Omega} + eV_E\} +$$
$$n(z_s/\mathrm{H}^+) \exp \frac{-1}{kT(\infty)} \{m_\mathrm{H} V_{G\Omega} + eV_E\}. \tag{33}$$

Neglecting m_e in comparison with m_O and m_H it may be seen that

$$n(z/\mathrm{O}^+) = n(z_s/e)^{\frac{1}{2}} \nu(z/\mathrm{O}^+) \Big/ \{\nu(z/\mathrm{O}^+) + \nu(z/\mathrm{H}^+)\}^{\frac{1}{2}}, \tag{34}$$

$$n(z/\mathrm{H}^+) = n(z_s/e)^{\frac{1}{2}} \nu(z/\mathrm{H}^+) \Big/ \{\nu(z/\mathrm{O}^+) + \nu(z/\mathrm{H}^+)\}^{\frac{1}{2}}, \tag{35}$$

and

$$n(z/e) = n(z_s/e)^{\frac{1}{2}} \{\nu(z/\mathrm{H}^+) + \nu(z/\mathrm{O}^+)\}^{\frac{1}{2}}, \tag{36}$$

where

$$\nu(z/\mathrm{O}^+) = n(z_s/\mathrm{O}^+) \exp \frac{-1}{kT(\infty)} \{m_\mathrm{O} V_{G\Omega}\}, \tag{37}$$

and

$$\nu(z/\mathrm{H}^+) = n(z_s/\mathrm{H}^+) \exp \frac{-1}{kT(\infty)} \{m_\mathrm{H} V_{G\Omega}\}. \tag{38}$$

Except that they include allowance for the rotation of the Earth $\nu(z/\mathrm{O}^+)$ and $\nu(z/\mathrm{H}^+)$ are what the number densities of O^+ and H^+ ions would be if the distribution of each species of ion beyond the reference level were the same as the distribution of the parent atom (neglecting any effect of escape).

Several points of interest emerge.

(*a*) The diffusion distributions (34) and (35) are such that if the condition (27) for charge transfer equilibrium is satisfied at the reference level, z_s, it is closely satisfied well above this level. Indeed, deviations only occur because of the influence of the rotation of the Earth on the distributions of ions and the influence of escape on the distribution of hydrogen atoms.

(*b*) In the thermosphere and lower exosphere

$$\nu(z/\mathrm{O}^+) \gg \nu(z/\mathrm{H}^+). \tag{39}$$

It follows from (34) that in this region

$$n(z/\mathrm{O}^+) \simeq n(z_s/\mathrm{O}^+) \exp \frac{-1}{kT(\infty)} \left\{\frac{m_\mathrm{O} V_{G\Omega}}{2}\right\} \tag{40}$$

which implies that the scale height of the O^+ ions is very closely twice that of the O atoms. The result is well known. For the same region (35) gives

$$n(z/\mathrm{H}^+) = \nu(z/\mathrm{H}^+) \exp \frac{1}{kT(\infty)} \left\{\frac{m_\mathrm{O} V_{G\Omega}}{2}\right\} \tag{41}$$

signifying that the number density of the H^+ ions is increasing with altitude at almost the same rate (per unit number density) as the number density of O^+ ions is decreasing. This was first recognised by Dungey[3]. It has not been taken into account in some recent investigations.

(*c*) Comparison of (34) and (35) shows that the altitude at which O^+ ions and H^+ ions are

equally abundant is determined by

$$\exp\frac{-1}{kT(\infty)}\{(m_{\mathrm{O}}-m_{\mathrm{H}})V_{G\Omega}\}=\frac{n(z_s/\mathrm{H}^+)}{n(z_s/\mathrm{O}^+)}=\frac{n(z_s/\mathrm{H})}{\sigma n(z_s/\mathrm{O})}\,. \tag{42}$$

At this altitude z_s the number-density of electrons is

$$n(z_s/e)=fn(z_s/\mathrm{O}^+)\ \ (=fn(z_s/e)) \tag{43}$$

with

$$f=2^{\frac{1}{2}}\left\{\frac{n(z_s/\mathrm{H})}{\sigma n(z_s/\mathrm{O})}\right\}^{8/15}\,. \tag{44}$$

(*d*) Sufficiently high in the exosphere

$$\nu(z/\mathrm{H}^+)\gg\nu(z/\mathrm{O}^+) \tag{45}$$

so that from (34), (35) and (36) it follows that

$$n(z/\mathrm{O}^+)\simeq An(z_s/\mathrm{O}^+)\exp\frac{-1}{kT(\infty)}\times\{(m_{\mathrm{O}}-\tfrac{1}{2}m_{\mathrm{H}})V_{G\Omega}\} \tag{46}$$

and

$$n(z/e)=n(z/\mathrm{H}^+)\simeq An(z_s/\mathrm{H}^+)\exp\frac{-1}{kT(\infty)}\times\left\{\frac{m_{\mathrm{H}}}{2}V_{G\Omega}\right\} \tag{47}$$

with

$$A=\left[\frac{n(z_s/\mathrm{O}^+)}{n(z_s/\mathrm{H}^+)}\right]^{\frac{1}{2}}=\left[\frac{\sigma n(z_s/\mathrm{O})}{n(z_s/\mathrm{H})}\right]^{\frac{1}{2}}\,. \tag{48}$$

In the important region where (45) is satisfied and where in addition

$$V_\Omega\ll V_G \tag{49}$$

the scale height for O^+ ions is therefore almost the same as that for O atoms (which is of relevance in connexion with calculations on the intensity-distribution of the Van Allen radiation (cf. Singer[23]); and the scale height for H^+ ions is twice that for H atoms (in hydrostatic equilibrium).

The number densities of the ions at very great altitudes are A times greater than they would be if the asymptotic forms of the distributions were valid throughout the region above the standard level. In a sense A, which is a large number, may be regarded as an amplification factor.

Noting that

$$n(z/\mathrm{O})=n(z_s/\mathrm{O})\exp\frac{-1}{kT(\infty)}\{m_{\mathrm{O}}V_G\} \tag{50}$$

it is apparent from (46) that if (45) holds then

$$n(z/\mathrm{O}^+)\Big/n(z/\mathrm{O})=\frac{\sigma^{\frac{1}{2}}n(z_s/\mathrm{O}^+)}{[n(z_s/\mathrm{O})\,n(z_s/\mathrm{H})]^{\frac{1}{2}}}\times\exp\frac{-1}{kT(\infty)}\{m_{\mathrm{O}}V_\Omega-\tfrac{1}{2}m_{\mathrm{H}}V_G\}. \tag{51}$$

If it were not for the effect of escape the corresponding relation for hydrogen would be

$$n(z/\mathrm{H}^+)\Big/n(z/\mathrm{H})=\frac{n(z_s/\mathrm{O}^+)}{[\sigma n(z_s/\mathrm{O})\,n(z_s/\mathrm{H})]^{\frac{1}{2}}}\exp\frac{-1}{kT(\infty)}\left\{\frac{m_{\mathrm{H}}}{2}(V_\Omega-V_G)\right\}\,. \tag{52}$$

The information available on the number densities concerned indicates that the multiplying factors on the right of (51) and (52) are of the order of unity.

4.3. *Results*

Since the distributions of the ions depend on several parameters it is scarcely feasible to present results covering every case of interest. Fortunately, the computations in any particular case are light. We shall confine ourselves to treating a single representative case and to describing the general trends.

The following values were assigned to the various parameters:

z_s (of Section 4.2) $=z_s$ (of Section 3.1) $=500$ km

$T(\infty)=1250\,^\circ\mathrm{K}$

$n(500\ \mathrm{km}/\mathrm{O})=2{\cdot}5\times10^7/\mathrm{cm}^3$ (Table 1)

$n(500\ \mathrm{km}/\mathrm{H})=1{\cdot}0\times10^4/\mathrm{cm}^3$ (Section 3.2)

and

$$n(500\ \mathrm{km}/\mathrm{O}^+)=4{\cdot}0\times10^5/\mathrm{cm}^3$$

Substitution of the adopted number densities in (27) yields

$$n(500\ \mathrm{km}/\mathrm{H}^+)=1{\cdot}8\times10^2/\mathrm{cm}^3$$

approximately. The distributions of ions and electrons in the altitude-range from 500 km to 2000 km were obtained from (34), (35) and (36) and are depicted in Fig. 6 on a logarithmic scale. It is seen that $n(z/\mathrm{O}^+)$ falls off sharply

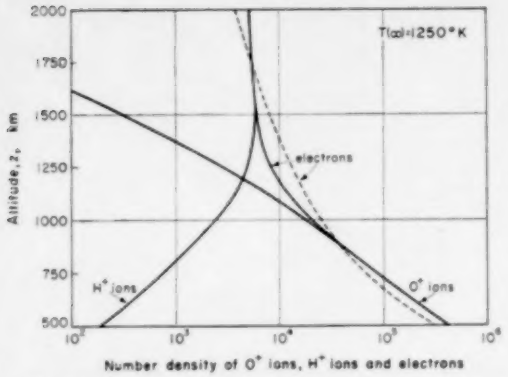


Fig. 6. Distributions of O^+ ions, H^+ ions and electrons :———Calculated (representative case described in text); — — — NRL model (arbitrary units).

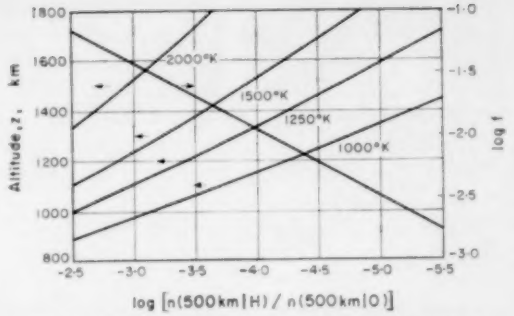


Fig. 7. The group of four curves with $T(\infty)$ marked on each give altitude of level where O^+ and H^+ ions are equally abundant (left vertical scale). The isolated curve—which is independent of $T(\infty)$ —gives the factor C by which the electron density decreases in going from the 500 km level to this level (right vertical scale).

and that the gradient increases near the altitude z_e at which the $n(z/O^+)$ and $n(z/H^+)$ curves cross; that $n(z/H^+)$ rises steeply to reach a broad maximum somewhat beyond z_e; and that $n(z/e)$ decreases rapidly up to about z_e but decreases very slowly above this altitude, in fair over-all agreement with the form of the Naval Research Laboratory model of Jackson which is based on a compilation of rocket, satellite and whistler data (cf. Newell[24]). In the region where $n(z/O^+)$ is much less than $n(z/H^+)$ the asymptotic formulae (46) and (47) may be used. Since the amplification factor A, defined in (48), is here about 47 these formulae may be written

$$n(z/O^+)=1{\cdot}9\times10^7\exp[-15{\cdot}5m_H V_{G\Omega}/1250k] \quad (53)$$

and

$$n(z/H^+)=8{\cdot}5\times10^3\exp[-0{\cdot}5m_H V_{G\Omega}/1250k]. \quad (54)$$

The numerical value of $m_H V_{G\Omega}/1250k$ may be read from Fig. 5. For comparison, it may be noted that the corresponding distributions of neutral atoms are given by

$$n(z/O)=2{\cdot}5\times10^7\exp[-16m_H V_G/1250k] \quad (55)$$

and (neglecting the effect of escape) by

$$n(z/H)=1{\cdot}0\times10^4\exp[-m_H V_G/1250k]. \quad (56)$$

Using (42) and (43) computations were carried out on the altitude z_e at which the number densities of O^+ and H^+ ions are the same and on the factor f by which the electron density falls in going from reference altitude z_s (500 km) to this altitude. Fig. 7 shows the results. It may be noted that with the abundance of atmospheric hydrogen finally adopted in Section 3.2 the independent variable $n(500\text{ km/H})\Big/n(500\text{ km/O})$ is 2×10^{-3}, 4×10^{-4}, 8×10^{-5} or 2×10^{-5} according as $T(\infty)$ is 1000°K, 1250°K, 1500°K or 2000°K.

According to the theory there is a marked change in the slope of an $n(z/e)$ curve near altitude z_e (cf. Fig. 6). Such a feature is not apparent in the Naval Research Laboratory model (cf. Fig. 6) presumably because the model is composite in character.

The $n(z/e)$ curve obtained by Berning[25] from measurements made from a rocket flown in the early morning (when the exosphere is relatively cool) gives some support to the theory. This curve exhibits a marked change of shape at a rather ill-defined altitude of about 1100 km. Tentatively taking z_e to be 1100 km it may be seen from Fig. 7 that the values of $n(500\text{ km/H})\Big/n(500\text{ km/O})$ corresponding to temperatures $T(\infty)$, of 1000°K and 1250°K are 2×10^{-4} and 1×10^{-3} respectively. These values, being sensitive to the assumed z_e, may be considerably in error.

Berning's measurements give the factor f to be 0·1 or rather greater. Referring again to Fig. 7 it might be thought that a value of $n(500\,\text{km/H})/n(500\,\text{km/O})$ as high as 5×10^{-3} is implied. However, it is unlikely that the distribution of the electrons with altitude at the time of the flight was the steady-state distribution.

Following a proposal made by Storey[26] regarding the detection of the protons in the exosphere, Barrington and Nishizaki[27] have recently succeeded in deriving information on z_e from two whistler dispersion curves. One of these was recorded late in an afternoon of December 1957, and the other was recorded early in a morning of March 1958. The electron densities were different. Rather surprisingly the results obtained are in close agreement. Both curves indicate z_e to be about 1000 km. The corresponding value of

$$n(500\,\text{km/H})\Big/n(500\,\text{km/O})$$

is 7×10^{-4} if $T(\infty)$ is 1000°K, is 3×10^{-3} if $T(\infty)$ is 1250°K, is 8×10^{-3} if $T(\infty)$ is 1500°K and is absurdly high if $T(\infty)$ is 2000°K (Fig. 7).

Summarizing the preceding few paragraphs the observed distributions of charged particles suggest that the hydrogen content of the exosphere is considerable (perhaps with

$$n(500\,\text{km/H})\Big/n(500\,\text{km/O})$$

about 1×10^{-3} or even greater) and suggest also that the temperature of the exosphere is in general moderate ($T(\infty)$ about 1250°K or less). The evidence, however, is inconclusive.

Fig. 8 shows the asymtotic form (47) of the steady-state electron density distribution in the equatorial plane*. A semi-logarithmic scale is used and the independent variable is chosen to be r of (32) that is, to be the distance from the centre of the Earth expressed in Earth-radii.

The predicted absolute electron densities may be obtained by multiplying the relative

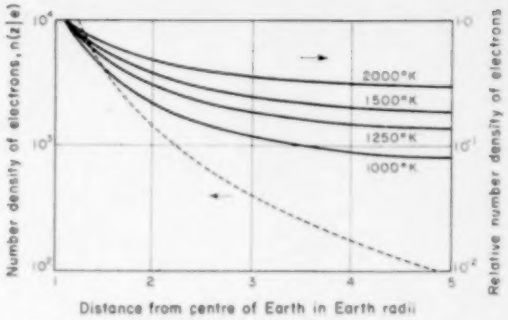


Fig. 8. Electron density distributions
—Calculated $T(\infty)$ being as indicated (right vertical scale—which is relative but may be made absolute as described in text);
— — Average of curves obtained by Smith and Helliwell from whistlers (left vertical scale).

electron densities depicted in Fig. 8 by

$$C = n(500\,\text{km/O}^+)\times\{n(500\,\text{km/H})/n(500\,\text{km/O})\}^{\frac{1}{2}}. \qquad (57)$$

Rather arbitrarily taking $n(500\,\text{km/O}^+)$ to be $4\times10^5/\text{cm}^3$, taking $n(500\,\text{km/H})$ to be as indicated at the end of Section 3.2 and taking $n(500\,\text{km/O})$ to be as in Table 1, it is found that when $T(\infty)$ is 1000°K, 1250°K, 1500°K and 2000°K, C is $1{\cdot}9\times10^4/\text{cm}^3$, $8{\cdot}0\times10^3/\text{cm}^3$, $3{\cdot}5\times10^3/\text{cm}^3$ and $1{\cdot}6\times10^3/\text{cm}^3$ respectively.

It is seen from Fig. 8 that the electron density distribution given by the theory has a much smaller gradient at great distances than has the electron density distribution derived from nose whistlers by Smith and Helliwell[28] (who drew attention to the discrepancy). An obvious defect of the theory in the simple form presented is that it does not take into account that conditions in the upper part of the F-layer depend on the latitude. This dependence influences the electron density distribution in the equatorial plane because the diffusion of charged particles in the exosphere is appreciable only along the geomagnetic lines. However, quite a narrow latitude range corresponds to a wide altitude range: thus the latitudes corresponding to the levels at 2, 3, 4 and 5 Earth radii are 45°, 55°, 60° and 63° respectively (cf. Chapman and Suguira[29]). To account for the discrepancy apparent in Fig. 8 some relevant property of the F-layer would have to change remarkably rapidly with latitude. It is probable

* The asymptotic form is followed when condition (45) is satisfied as it is in almost all of the region covered by the curves drawn.

that another explanation must be sought. The inclusion of an r^{-3} multiplying factor in the formula from which the distribution is computed brings about good accord with whistler data (Pope[30]); but no acceptable theoretical justification for the inclusion of this factor has yet been given. Dungey[31] has recently suggested that the equilibrium under the gravitational and electrostatic forces may be upset by a mixing of the interplanetary plasma into the outer atmosphere. The hypothesis has not been developed quantitatively.

APPENDIX

Possible intermediate zone

In the steady state equation (27) will be correct at all altitudes unless an intermediate zone exists in which there is neither charge transfer equilibrium nor complete diffusive equilibrium. Such an intermediate zone exists only if the altitude $z_{d\mathrm{H}^+}$ of the level at which the distribution of H^+ ions ceases to be controlled by charge transfer and becomes controlled by diffusion is less than the altitude $z_{d\mathrm{O}^+}$ of the level at which the distribution of O^+ ions ceases to be controlled by photoionization and recombination processes and becomes controlled by diffusion.

The value of $z_{d\mathrm{O}^+}$ is certainly not greater than about 500 km and is probably rather less (cf. Ferraro[32], Johnson[33]). Over a moderate altitude interval within the region $z < z_{d\mathrm{O}^+}$ it is a sufficient approximation to take

$$n(z/\mathrm{O}^+) \propto n(z/\mathrm{O})^{(1/2)-\delta} \qquad \text{(A.1)}$$

where δ (which is a measure of the extent to which the distribution of O^+ ions departs from diffusive equilibrium) is less than $\frac{1}{2}$ and is positive. Noting that $n(z/\mathrm{H})$ is effectively constant throughout the region of interest it may readily be proved that

$$n(z/\mathrm{H}^+) \propto n(z/\mathrm{O})^{-(1/2)-\delta} \qquad \text{(A.2)}$$

for

$$z < z_{d\mathrm{H}^+}; \qquad \text{(A.3)}$$

and that

$$n(z/\mathrm{H}^+) \propto n(z/\mathrm{O})^{(-1/2)+\delta} \qquad \text{(A.4)}$$

for

$$z_{d\mathrm{H}^+} < z < z_{d\mathrm{O}^+} \qquad \text{(A.5)}$$

The values of $n(z/\mathrm{H}^+)\Big/n(z/\mathrm{H})$ given in the main text satisfy the charge transfer equilibrium equation (27) even in the region of complete diffusion equilibrium

$$z > z_{d\mathrm{H}^+} \text{ and } z_{d\mathrm{O}^+}. \qquad \text{(A.6)}$$

Should the intermediate zone defined by (A.5) exist they would clearly be too high by the factor

$$q = \left\{\frac{n(z_{d\mathrm{H}^+}/\mathrm{O})}{n(z_{d\mathrm{O}^+}/\mathrm{O})}\right\}^{2\delta} = \frac{n(z_{d\mathrm{H}^+}/\mathrm{O})}{n(z_{d\mathrm{O}^+}/\mathrm{O})}\,\frac{n(z_{d\mathrm{O}^+}/\mathrm{O}^+)^2}{n(z_{d\mathrm{H}^+}/\mathrm{O}^+)^2} \qquad \text{(A.7)}$$

It is unlikely that q is large. Thus taking what will later be seen to be an extreme case suppose that $z_{d\mathrm{H}^+}$ is 300 km and that $z_{d\mathrm{O}^+}$ is 500 km. On the 1500°K model atmosphere

$$\frac{n(300\text{ km}/\mathrm{O})}{n(500\text{ km}/\mathrm{O})} = 9{\cdot}7 \qquad \text{(A.8)}$$

The available measurements (cf. Nisbet[34]) indicate

$$0{\cdot}2 < \frac{n(500\text{ km}/\mathrm{O}^+)}{n(300\text{ km}/\mathrm{O}^+)} < 0{\cdot}6. \qquad \text{(A.9)}$$

Hence even in this case the maximum value of q is only about 3·5 and the mean value is little above unity.

We shall now consider what must be the magnitude of the rate coefficient K of (26) if there is to be charge transfer equilibrium, confining our attention to high latitudes where diffusion is not retarded by the geomagnetic field.

For charge transfer equilibrium it is necessary that

$$Kn(z/\mathrm{H}^+)\,n(z/\mathrm{O}) \gg \frac{-dS(z/\mathrm{H}^+)}{dz} \qquad \text{(A.10)}$$

$-S(z/\mathrm{H}^+)$ being the downward flux of H^+ ions when the steady state is reached. In an isothermal region

$$-S(z/\mathrm{H}^+) = D(z/\mathrm{H}^+)\left\{\frac{dn(z/\mathrm{H}^+)}{dz} + \frac{n(z/\mathrm{H}^+)}{H(z/\mathrm{H}^+)}\right\} \qquad \text{(A.11)}$$

where $D(z/\mathrm{H}^+)$ is the diffusion coefficient and $H(z/\mathrm{H}^+)$ is the scale height appropriate to

diffusion equilibrium. The rate of diffusion of H^+ ions in the upper atmosphere is determined by collisions with O^+ ions rather than with O atoms so that

$$D(z/H^+) = b \Big/ n(z/O^+) \qquad (A.12)$$

where b is a constant. Using (A.1) and (A.2) and neglecting certain small terms it may hence be shown that (A.10) is equivalent to

$$K \gg b \Big/ 2n(z/O^+)\, n(z/O)\, H(z/O)^2 \qquad (A.13)$$

According to the standard formula cited by Chapman and Cowling[35] b is $8{\cdot}4 \times 10^{15}$ cm^{-1} sec^{-1} at 1500°K. Taking

$$\begin{aligned} n(300\text{ km}/O^+) &= 1 \times 10^6/cm^3; \\ n(300\text{ km}/O) &= 4{\cdot}2 \times 10^8/cm^3; \\ H(300\text{ km}/O) &= 88\text{ km} \end{aligned} \qquad (A.14)$$

and

$$\begin{aligned} n(500\text{ km}/O^+) &= 5 \times 10^5/cm^3; \\ n(500\text{ km}/O) &= 4{\cdot}2 \times 10^7/cm^3; \\ H(500\text{ km}/O) &= 93\text{ km}. \end{aligned} \qquad (A.15)$$

it is found from (A.13) that in the high latitude case there will be charge transfer equilibrium up to an altitude of 300 km if K is much greater than 1×10^{-13} cm^3/sec and up to an altitude of 500 km if K is much greater than 2×10^{-12} cm^3/sec. The values of K needed to ensure charge transfer equilibrium up to the same altitudes at low latitudes are of course smaller.

The rate coefficient describing close encounters between ions and atoms of polarizability P is given by

$$k = 2\pi (e^2 P/\mu)^{\frac{1}{2}} \qquad (A.16)$$

where μ is the reduced mass (Gioumousis and Stevenson[36]). For oxygen P is $7{\cdot}7 \times 10^{-25}$ cm^3 (Alpher and White[37]). Substitution in (A.16) yield that for close $H^+ - O$ encounters

$$k = 2 \times 10^{-9}\ cm^3/sec \qquad (A.17)$$

Hence the values of K given in the preceding paragraph correspond to values of only 5×10^{-5} and 1×10^{-3} respectively for the probability of charge transfer per close collision.

Acknowledgement—The research reported has been sponsored by the Air Force Cambridge Research Center, Geophysics Research Directorate of the Air Research and Development Command U.S. Air Force, through its European Office and Contract No. AE 61(052)-131.

REFERENCES

1. C. O. Hines, *J. Geophys. Res.* **65,** 2563 (1960).
2. D. R. Bates, *Proc. Roy. Soc.* **A253,** 451 (1959).
3. J. W. Dungey, *The Physics of the Ionosphere* p. 229. Report of the 1954 Cambridge Conference. Physical Society, London (1955).
4. P. Mange, *J. Geophys. Res.* **65,** 3833 (1960).
5. D. G. King-Hele and D. M. C. Walker, *Nature, Lond.* **186,** 928 (1960).
6. M. Nicolet, *Space Research* (Edited by H. Kallmann Bijl) p. 45. North Holland Publishing Co., Amsterdam (1960).
7. M. Nicolet, *Planet. Space Sci.* **5,** 1 (1961).
8. S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-uniform Gases* (2nd ed.) p. 244. Cambridge University Press (1952).
9. L. Spitzer, *Atmospheres of the Earth and Planets* (Edited by G. P. Kuiper) p. 213. University Press, Chicago (1949).
10. K. E. Grew and T. L. Ibbs, *Thermal Diffusion in Gases* p. 28. Cambridge University Press (1952).
11. E. J. Opik and S. F. Singer, *Phys. Fluids* **2,** 653 (1959).
12. S. F. Singer, *Planet. Space Sci.* **2,** 165 and 263 (1960).
13. D. R. Bates and M. R. C. McDowell, *J. Atmos. Terr. Phys.* **11,** 200 (1957); **16,** 393 (1959).
14. D. R. Bates and M. Nicolet, *J. Geophys. Res.* **55,** 301 (1950).
15. J. E. Kupperian, E. T. Byram, T. A. Chubb and H. Friedman, *Ann. Geophys.* **14,** 329 (1958); *Planet. Space Sci.* **1,** 7 (1959).
16. J. D. Brandt and J. W. Chamberlain, *Astrophys. J.* **130,** 670 (1959).
17. O. Struve, *Sky Telesc.* **17,** 445 (1958).
18. I. S. Shklovsky, *Planet. Space Sci.* **1,** 63 (1959).
19. F. S. Johnson and R. A. Fish, *Astrophys. J.* **131,** 502 (1960).
20. J. D. Purcell and R. Tousey, *J. Geophys. Res.* **65,** 370 (1960).
21. D. R. Bates and N. Lynn, *Proc. Roy. Soc.* **A253,** 141 (1959).
22. V. G. Istomin, C.S.A.G.I. Assembly, Moscow (1958).
23. S. F. Singer, *J. Geophys. Res.* **65,** 2577 (1960).
24. H. E. Newell *Physics of the Upper Atmosphere* (Edited by J. A. Ratcliffe) p. 73. Academic Press, New York and London (1960).
25. W. W. Berning, *J. Geophys. Res.* **65,** 2589 (1960).

26. L. R. O. Storey, *Canad. J. Phys.* **34,** 1153 (1956).
27. R. E. Barrington and I. Nishizaki, *Canad. J. Phys.* **38,** 1642 (1960).
28. R. L. Smith and R. A. Helliwell, *J. Geophys. Res.* **65,** 2583 (1960).
29. S. Chapman and M. Suguira, *J. Geophys. Res.* **61,** 485 (1956).
30. J. H. Pope, *J. Geophys. Res.* **66,** 67 (1961).
31. J. W. Dungey, *Phys. Rev. Lett.* **6,** 47 (1961).
32. V. C. A. Ferraro, *Terr. Magn. Atmos. Elect.* **50** 215 (1945).
33. F. S. Johnson, *J. Geophys. Res.* **65,** 577 (1960).
34. J. S. Nisbet, *J. Geophys. Res.* **65,** 2597 (1960).
35. S. Chapman and T. G. Cowling *The Mathematical Theory of Non-uniform Gases* (2nd ed.) p. 179. Cambridge University Press (1952).
36. G. Gioumousis and D. P. Stevenson, *J. Chem. Phys.* **29,** 294 (1958).
37. R. A. Alpher and D. R. White, *Phys. Fluids* **2,** 159 (1959).

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 274-282. Printed in Great Britain.

LE GUIDAGE DES WHISTLERS PAR LE CHAMP MAGNETIQUE

R. GENDRIN

Département C.D.S., Centre National d'Etudes des Télécommunications,
Issy-les-Moulineaux (Seine), France

(*Received* 10 *September* 1960)

Abstract—In this paper the question is examined of how the v.l.f. radio-waves are guided along the magnetic field. Energy passes through the magnetic field under two sets of conditions. Corresponding to the "nose-whistlers" explained by Helliwell, the first one occurs when the wave-normal itself is in the direction of the magnetic field. This does not happen in the second case when the remarkable property is also shown that all frequencies are propagated at the same velocity $V_0 = cf_H/2f_0$ (f_H gyrofrequency, f_0 frequency of the plasma). Considerations of energy point out that, if such a propagation is not easily observable in the case of an isotropic emission, it is not the same thing for an emission produced by Čerenkov effect, which is able to produce all energy by this mode of propagation, provided the particle's velocity has a low fixed value ($\sim$ 10,000 km/sec in the exosphere). All frequencies being emitted at the same time and following the same path wtih the same velocity, we can explain the broad-band noise observed during the reception of whistlers. The required velocity of particles is exactly the velocity V_0. This coincidence is explained in an appendix, and extended to other anisotropic media.

Résumé—On étudie le guidage des ondes de très basse fréquence par le champ magnétique. On montre que l' énergie peut se propager le long du champ magnétique dans deux cas. Le premier, qui correspond aux "nose-whistlers" interprétés par Helliwell, est celui dans lequel la normale à l'onde de phase est elle-même dirigée suivant le champ magnétique. Il n'en est pas de même dans le second cas qui possède par ailleurs la propriété remarquable que toutes les fréquences se propagent à la même vitesse $V_0 = cf_H/2f_0$ (f_H gyrofréquence, f_0 fréquence de plasma). Des considérations d'énergie montrent que si une telle propagation est difficilement observable dans le cas d'une émission isotrope, il n'en est pas de même dans le cas d'une émission produite par effet Čerenkov. On montre qu'au contraire il existe une vitesse des particules ($\sim$ 10,000 km/s dans l'exosphère) permettant d'émettre toutes les fréquences dans ce mode. On a ainsi une explication possible des bruits à large bande observés dans l'écoute des whistlers. La vitesse nécessaire des particules est justement la vitesse V_0. Une explication de cette coïncidence est donnée en appendice. Elle se généralise à d'autres milieux anisotropes.

1. INTRODUCTION

Le guidage des ondes électromagnétiques de basse fréquence—c'est-à-dire la propriété qu'a le rayon électromagnétique d'être peu incliné sur le champ magnétique—est connu depuis longtemps. Storey[1] est le premier qui montra que l'énergie de whistlers se propage dans l'exosphère, en faisant un angle faible avec le champ magnétique, même si la normale à l'onde forme avec lui un angle grand. Il montra également que, sous certaines conditions, cet angle restait inférieur à un angle limité de 19° 29′. Mais il ne s'était occupé que de fréquences faibles vis-à-vis de la gyrofréquence électronique. Helliwell donna une formule moins approchée—limitée au cas de la propagation le long du champ magnétique—et interpréta ainsi le phénomène des nose-whistlers[2], mais il ne se préoccupa pas entièrement du phénomène de guidage. Son attention fut plutôt dirigée vers un autre processus de guidage : le guidage par les fibres[3], dont nous ne nous occuperons pas ici.

Le travail que nous présentons a pour but d'établir la nature du guidage par le champ magnétique dans l'approximation de Helliwell,

non point tant pour "combler une lacune" que pour faire apparaître un phénomène nouveau, lié au mode de propagation des whistlers et susceptible d'expliquer certains bruits de très basse fréquence.

2. FORMULES RELATIVES À LA PROPAGATION

Données de base

On sait que[4], en l'absence de collisions, pour le mode extraordinaire et pour une propagation quesi-longitudinale, l'indice de réfraction de phase est donné par la formule*

$$n^2 = 1 - \frac{X}{1 - Y\cos\theta} \tag{1}$$

dans laquelle

$X = (f_0/f)^2$

$Y = f_H/f$

f_0, fréquence de plasma
$f_0^2 = Ne^2/4\pi^2 m\varepsilon_0$†

f_H, gyrofréquence électronique
$f_H = eH_0\mu_0/2\pi m$

θ angle de la normale à l'onde avec le champ magnétique directeur H_0

que nous récrirons sous la forme

$$n^2 = 1 + \frac{a^2}{x(\cos\theta - x)} \tag{2}$$

avec

$a = f_0/f_H$ paramètre caractéristique du milieu

$x = f/f_H$ variable réduite définissant la fréquence de l'onde électromagnétique employée.

On sait également que, dans un milieu anisotrope, le rayon se propage suivant une direction faisant un angle α avec la normale à l'onde‡, et avec une vitesse V vérifiant:

$$tg\alpha = \frac{1}{v_\varphi}\frac{\partial v_\varphi}{\partial\theta}, \tag{3}$$

$$V = v_g \sec\alpha \tag{4}$$

$$\frac{1}{v_g} = \frac{n'}{c} = \frac{1}{c}\frac{\partial(nx)}{\partial x} \tag{5}$$

v_φ vitesse de phase

v_g vitesse de groupe (au sens des milieux isotropes)

n' indice de groupe

V vitesse du rayon

La Fig. 1 rassemble ces définitions.

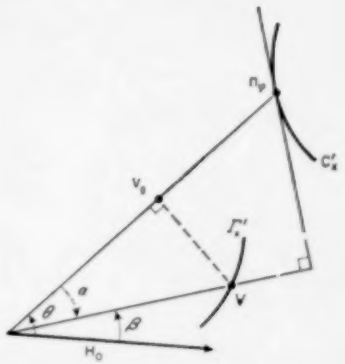


Fig. 1. Définitions. L'homothétie qui permet de passer des courbes C'_x et Γ'_x représentées sur cette figure, aux courbes C_x et Γ_x utilisées ultérieurement est définie dans le texte.

En reportant la valeur de l'indice de phase donnée par la formule (2) dans les équations (3) et (5) on obtient

$$n' = \frac{2x^3 - 4x^2\cos\theta + 2x\cos^2\theta + a^2\cos\theta}{2x^{1/2}[a^2 + x(\cos\theta - x)]^{1/2}(\cos\theta - x)^{3/2}} \tag{6}$$

$$tg\alpha = -\frac{a^2\sin\theta}{2(\cos\theta - x)[a^2 + x(\cos\theta - x)]} \tag{7}$$

Validité de l'approximation quasi-longitudinale (QL)

En l'absence de collisions, l'approximation QL est valable[4] tant que

$$\frac{1}{2}Y\frac{\sin^2\theta}{\cos\theta} \ll |1 - X|$$

ce qui se traduit, si $X > 1$, par

$$\frac{\sin^2\theta}{2\cos\theta} \ll \frac{a^2}{x} - x.$$

* Nous supposons dans ce qui suit $\cos\theta > 0$: Ceci ne change rien à la généralité mais nous alourdirait par la présence de barres de valeur absolue.

† Système MKSA rationalisé.

‡ La direction de propagation est également perpendiculaire à la tangente en M à la courbe d'indice.

Si l'on se place dans le cas le plus défavorable où $x=1$, on voit que l'expression (2) est exacte à mieux que 5 pour-cent près pour toute valeur de $\theta<\theta_0$. θ_0 est une fonction de a dont nous reproduisons quelques valeurs (Tableau 1).

Tableau 1. Valeurs limites de l'approximation quasi-longitudinale

a	1,4	1,8	2,3	2,6	3,2	4
θ_0	45°	60°	70°	75°	80°	85°

Nous voyons que, dans le cas usuel qui nous intéresse, a est grand (>3)* et *l'approximation QL est valable même pour de très grands angles.*

L'approximation de Helliwell

Elle consiste à négliger x devant a^2, ce qui est toujours le cas pour les whistlers (sauf dans la très basse ionosphère), puisque $x<1$ et $a^2\gg 1$.

Les formules (2), (6) et (7) deviennent alors:

$$n\sim\frac{a}{x^{1/2}(\cos\theta-x)^{1/2}} \tag{8}$$

$$n'\sim\frac{a\cos\theta}{2x^{1/2}(\cos\theta-x)^{3/2}} \tag{9}$$

$$tg\alpha\sim-\frac{\sin\theta}{2(\cos\theta-x)} \tag{10}$$

auxquelles s'ajoute

$$tg\beta\sim\frac{\sin\theta(\cos\theta-2x)}{1+\cos\theta(\cos\theta-2x)} \tag{11}$$

$\beta=\theta+\alpha$ étant l'angle de propagation du rayon avec le champ magnétique.

Courbes d'indice et de vitesse

Dans cette approximation nous avons construit pour six valeurs de x, et en nous limitant aux valeurs positives de θ, les courbes C_x (points de coordonnées polaires n/a, θ) et Γ_x (points de coordonnées polaires aV/c, β). Les courbes Φ_x (points de coordonnées polaires v_φ, θ) n'ont pas été construites, mais leur définition nous sera utile ultérieurement† (Fig. 2*a–f*).

* Sauf dans la très basse ionosphère.

† Elles sont inverses des courbes C_x dans l'inversion de pôle 0 et de puissance c/a.

Lorsque la courbe Γ_x présente un retour sur elle-même pour les valeurs positives de β (ce qui se produit pour tout $x<0{,}5$), la valeur maximum de β est inférieure à 19° 29′. Le tableau 2 donne les valeurs de β_m pour quelques valeurs de x, ainsique les valeurs de θ correspondantes.

Tableau 2. Valeur maximum de l'angle de propagation

x	0	0,05	0,1	0,2	0,3	0,4	0,5
θ	54°44′	51°12′	47°35′	40°02′	31°43′	21°41′	0°
β_m	19°29′	17°10′	14°52′	10°25′	6°12′	2°21′	0°

On remarque

(1) Que ces courbes ne ressemblent pas à celles données par Storey[1], même pour de faibles valeurs de x. Entre autres, et c'est normal, elles ne présentent pas de vitesse de propagation infinie. Cela tient à ce que nous n'avons pas négligé x devant $\cos\theta$. On remarque toutefois que la courbe $\Gamma_{0,025}$ présente une boucle très étroite qui s'apparente au point de rebroussement dessiné dans le travail original de Storey.

(2) Que la notion d'angle maximum n'a de sens que pour x faible, les angles β pouvant prendre d'importantes valeurs négatives. (Ce sont même les seules valeurs qu'ils prennent dès que $x>0{,}5$.)

(3) Que toutes les courbes Γ_x passent par le point ($\frac{1}{2}$, 0). En effet, lorsque $\cos\theta=2x$, $\beta=0$, d'après la formule (11); donc $\cos\alpha=2x$. Les formules (4) et (9) donnent alors $n'=a/x$ et

$$V=V_0=c/2a. \tag{12}$$

Il existe donc pour chaque fréquence un angle d'émission $\theta\neq 0$, tel que le rayon se propage strictement le long de la ligne de force magnétique, *avec une vitesse indépendante de la fréquence.*

(4) Que toutes les courbes C_x sont tangentes à une même droite perpendiculaire à H_0, d'équation $\rho=2/\cos\theta$.

Les deux phénomènes sont liés (voir appendice I).

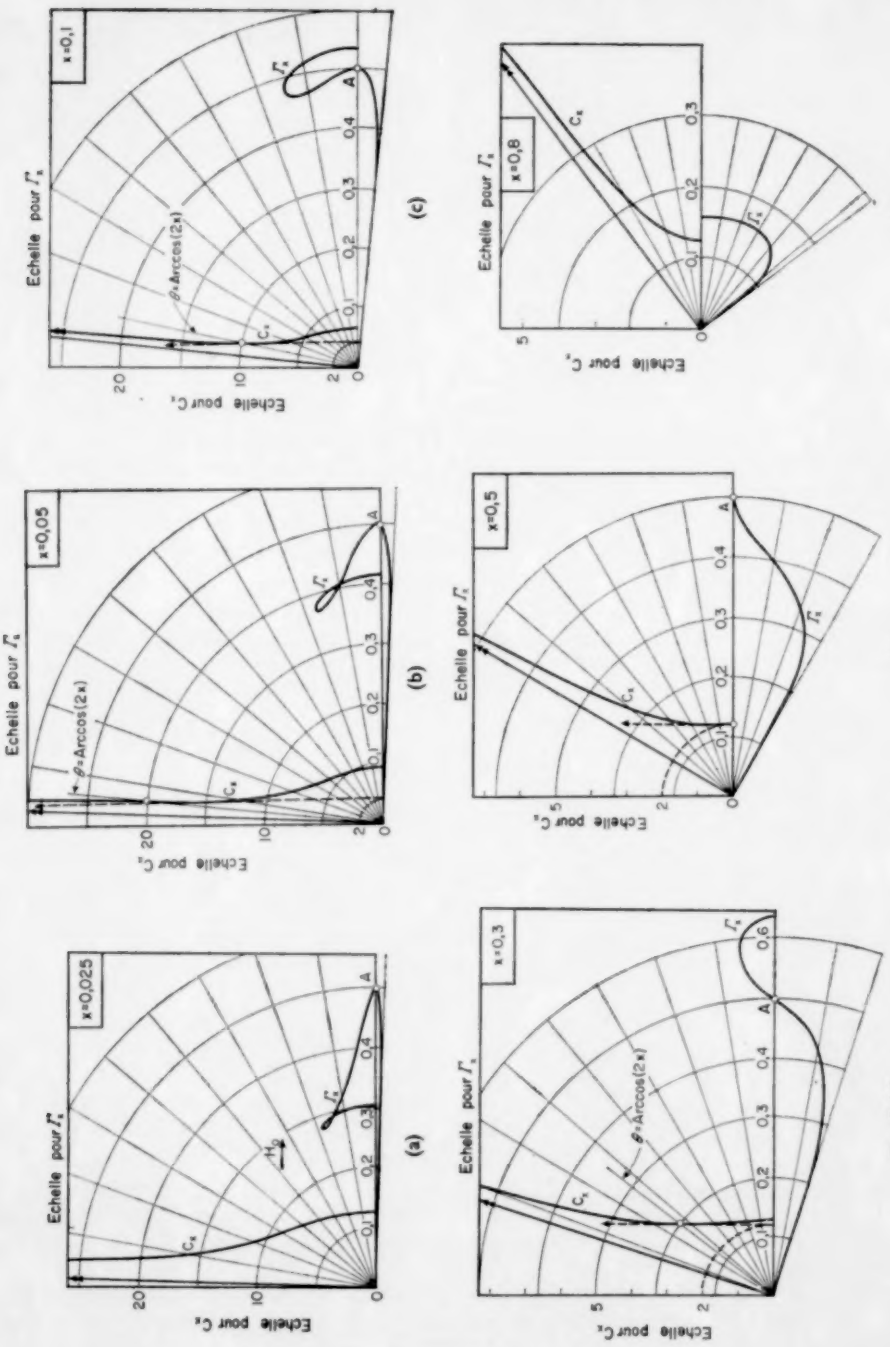


Fig. 2 (d)

Fig. 2 (a) Guidage pour $f/f_H = 0{,}025$; (b) Guidage pour $f/f_H = 0{,}05$; (c) Guidage pour $f/f_H = 0{,}1$; (d) Guidage pour $f/f_H = 0{,}3$; (e) Guidage pour la valeur limite $f/f_H = 0{,}5$; (f) Guidage pour $f/f_H > 0{,}5$. Les angles de propagation autorisés sont grands et les vitesses faibles. Les courbes Γ_x ne passent plus par le point A de coordonnées $\frac{1}{2}$, O.

3. PROPAGATION LE LONG DU CHAMP MAGNETIQUE

Les deux modes de propagation

L'équation (11) montre que, pour que $\beta=0$, il faut

$$\sin\theta=0$$
$$\text{ou} \qquad \cos\theta=2x \qquad (13)$$

Le premier cas correspond au nose-whistler de Helliwell; on a alors

$$n'=\frac{c}{V}=\frac{a}{2x^{1/2}(1-x)^{3/2}}. \qquad (14)$$

Dans l'hypothèse d'un milieu homogène (f_0 et f_H constant), on reçoit à l'autre extrémité de la ligne de force un signal ayant la forme indiquée Fig. 3.

Le second cas est celui qu'on a vu plus haut. La vitesse étant indépendante de la fréquence, le signal reçu à l'autre extrémité aura un spectre d'impulsion.

Si donc en un point, on émet toutes les fréquences, dans *toutes les directions*, et que l'on écoute à l'autre extrémité, on doit recevoir un signal dont le spectre fréquence en fonction du temps a l'aspect *complet* de la Fig. 3.

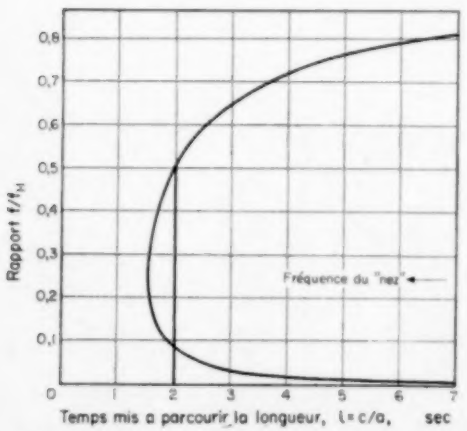


Fig. 3. Spectre complet obtenu pour une émission isotrope de toutes les fréquences. L'observation est effectuée à l'autre extrémité de la ligne de force. f_0 et f_H sont supposés constants.

Considérations d'énergie

Supposons la source précédente isotrope. Quelle est l'énergie que l'on peut attendre par le mode II dans un intervalle de temps 2 dt? On reçoit tout ce qui a été émis suivant les directions vérifiant la formule (13) et dont la vitesse V vérifie

$$c/2\,a-l/dt<V<c/2\,a+l/dt \qquad (15)$$

l étant la distance parcourue le long de H_0.

Etant donné l'allure des courbes Γ_x, la condition (15) correspond à des valeurs de $d\beta$ et de $d\theta$ très petites.

Par contre il est facile de voir que dans le premier mode on reçoit tout ce qui a été émis dans un angle $d\theta$ grand.

On a donc beaucoup plus d'énergie dans le premier mode que dans le second. Ce qui explique, qu'à notre connaissance, on n'ait pas enregistré de whistler ayant l'aspect de la Fig. 3.

On voit donc que par une émission isotrope —qui n'a pas de réalité physique dans le cas des whistlers causés par des éclairs*—on a peu de chances d'observer le mode II. Nous allons voir qu'un phénomène naturel, très important dans l'exosphère est *capable de produire toute l'énergie dans le mode II*.

4. APPLICATION À L'EFFET ČERENKOV

L'effet Cerenkov dans l'exosphère

L'effet Čerenkov produit par des particules d'origine solaire, peut être une cause importante du bruit de très basse fréquence observé dans l'écoute des whistlers. L'indice de réfraction étant en effet très élevé, il suffit de vitesses faibles par rapport à la vitesse de la lumière pour l'engendrer ($\sim$ 10,000 km/s). Ce phénomène a, entre autres, été mentionné par Ellis[5]. Gallet[6,7] en a étudié un cas particulier; celui où la vitesse de la particule est rigoureusement égale à la vitesse de phase. On a alors un mécanisme d'oscillation analogue à celui des

* À cause de la forte réfraction de l'ionosphère qui rend les angles θ très voisins à leur entrée dans l'exosphère.

(To face page 278)

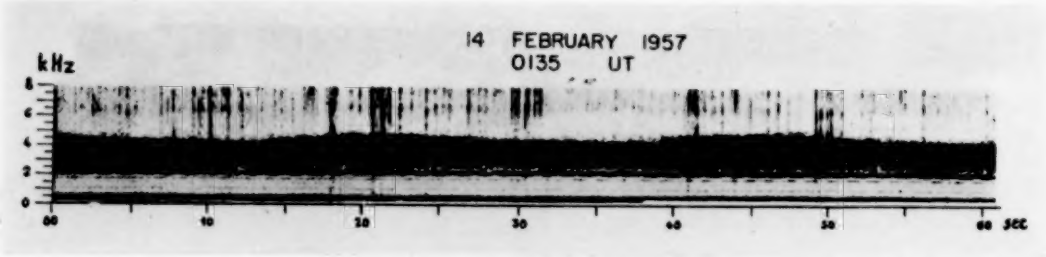


Fig. 4. Exemple de bruit à large bande observé dans l'écoute des whistlers. (Fig. 7 de la référence (6).

tubes à ondes progressives*. L'étude générale de l'effet Čerenkov dans un plasma a été mentionnée[9]. Nous renvoyons à l'appendice II pour l'étude géométrique, et nous allons tout de suite étudier le cas des whistlers vérifiant l'approximation de Helliwell.

Cas de l'approximation de Helliwell

Une particule chargée, de vitesse $w=\gamma c$, que nous supposons dirigée suivant H_0, émet la fréquence f dans un angle θ tel que

$$v_\varphi = w\cos\theta \tag{16}$$

Etant donné l'anisotropie du milieu, il y aura émission de fréquences différentes dans chaque angle θ. Calculons les angles d'émission de la fréquence $f=x f_H$. Ils vérifient

$$\cos\theta = \frac{v_\varphi}{w} = \frac{1}{\gamma n} = \frac{\sqrt{[x(\cos\theta - x)]}}{\gamma a}$$

d'où l'on tire

$$\gamma^2 a^2 \cos^2\theta - x\cos\theta + x^2 = 0 \tag{17}$$

équation qui montre que, pour un angle θ donné on a émission de deux fréquences. De même pour une fréquence donnée on a deux angles d'émission donnés par

$$\cos\theta = \frac{x}{2\gamma^2 a^2}\left[1 \pm \sqrt{(1-4\gamma^2 a^2)}\right]. \tag{18}$$

Ceci est vrai sous réserve que $4\gamma^2 a^2 < 1$, c'est-à-dire pour des vitesses w faibles; résultat déjà mis en évidence par Gallet[7].

Lorsque $\gamma = 1/2a$, les deux angles sont confondus, et leur valeur commune est donnée par

$$\cos\theta = 2x.$$

Cette valeur est justement celle pour laquelle le rayon se propage suivant H_0 avec une vitesse $V = c/2a$ indépendante de la fréquence.

Lorsque la condition

$$\gamma = 1/2a \tag{19}$$

est vérifiée, la propagation de toutes les fréquences émises par le passage de la particule se fait suivant le même trajet et à la même vitesse. On a là une interprétation possible des bandes de bruit observées entre 2 et 6 kHz et qui seraient ainsi une succession d'impulsions émises en même temps et arrivant en même temps.[10] La Fig. 4 en représente un exemple.

Il faut noter en plus que la vitesse de propagation commune des ondes est alors égale à celle de la particule:

$$w = V_0 \tag{20}$$

Nous donnons une explication de cette coincidence dans l'appendice II.

5. CONCLUSION

Par l'étude détaillée de la propagation des whistlers dans l'approximation de Helliwell, nous avons montré que chaque fréquence peut se propager le long du champ magnétique avec une vitesse V_0 ne dépendant que des propriétés du milieu. Dans le cas d'une émission isotrope, l'énergie reçue à l'autre extrémité de la ligne de force est trop faible pour que le phénomène puisse être observé. Par contre, l'effet Čerenkov, peut, pour une vitesse donnée des particules, émettre sélectivement chaque fréquence dans l'angle nécessaire pour que toute l'énergie se propage dans ce mode. La vitesse de la particule est alors égale à la vitesse de propagation "du rayon" commune à toutes les fréquences. On a là le pendant du mécanisme proposé par Gallet (vitesse de la particule = vitesse de phase) pour expliquer certains bruits v.l.f. Toutefois, ce mécanisme ne prend sa pleine puissance que lorsque f_0 et f_H sont constants. On peut penser avoir là un phénomène intéressant pour déterminer la densité électronique d'un plasma placé dans un champ magnétique connu, puisque c'est seulement pour la vitesse w vérifiant

$$w/c = f_H/2f_0*$$

que l'on a ce phénomène.

APPENDICE I

Quelques propriétés des courbes C_x dans le cas de l'approximation de Helliwell.

Nous allons montrer que la possibilité de propagation de toutes les fréquences le long de H_0 à la même

* L'équivalence avec l'effet Cerenkov en a d'ailleurs été montrée[8].

* Une valeur très proche de celle-ci joue également un rôle de valeur limite dans le mécanisme proposé par Gallet.

vitesse découle directement du fait que, dans l'approximation de Helliwell, *les courbes d'indice admettent pour enveloppe une droite perpendiculaire à* H_0.

Les courbes C_x forment en effet une famille de courbes à un paramètre (x) dont on obtient l'enveloppe en éliminant celui-ci entre les équations

$$n/a=[x(\cos\theta-x)]^{-1/2}$$

$$\frac{\partial(n/a)}{\partial x}=0 \qquad (21)$$

La relation (21) donne $\cos\theta=2x$ qui définit le point de contact M avec l'enveloppe et l'équation de celle-ci est

$$\rho=n/a=2/\cos\theta.$$

C'est une droite D perpendiculaire à H_0 qui coupe $0x$ en un point I tel que

$$0I=2.$$

Pour une onde dont la normale est dirigée suivant $0M$, le rayon se propage suivant $0x$, puisque $0x$ est alors perpendiculaire à la tangente en M à la courbe des indices.

Le point P, inverse de M dans l'inversion de pôle 0 et de puissance c/a est un point de la courbe Φ_x (Fig. 5). Mais puisqu'en M

$$\frac{\partial n}{\partial x}=0 \text{ on a : } v_g=v_\varphi$$

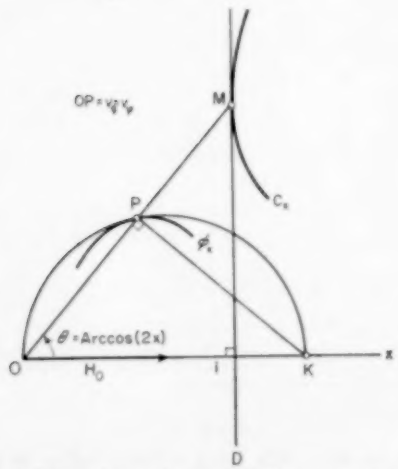


Fig. 5. L'enveloppe des courbes C_x est une droite D. Au point de contact avec l'enveloppe $\partial n/\partial x=0$ donc $v_g=v_\varphi$.

et par conséquent la vitesse de propagation du rayon est mesurée par $0K$; K étant à l'intersection de $0x$ et de la perpendiculaire menée en P à $0P$. Mais P décrit lorsque x varie le cercle inverse de D, et K est donc l'inverse de I. Il ne dépend pas de la fréquence.

K *est un point fixe* et l'on a

$$0K=V_0=c/2a.$$

APPENDICE II

Remarques géométriques sur l'effet Cerenkov dans un milieu anisotrope et dispersif.

Soit une particule se déplaçant à la vitesse w suivant une direction faisant un angle β avec une direction $0x$ de référence (Fig. 6). Quels sont les angles θ (définissant la normale à l'onde) suivant lesquels j'engendre une fréquence f déterminée? Ceux pour lesquels

$$v_\varphi=w\cos(\theta-\beta).$$

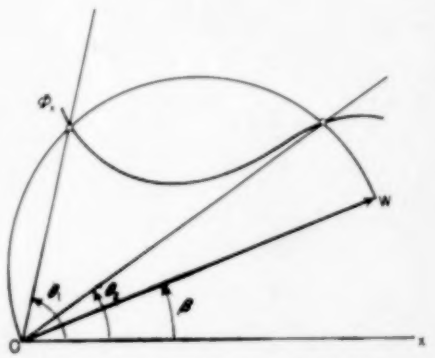


Fig. 6. Effet Čerenkov dans un milieu anistrope. La vitesse de la particule est représentée par 0W. L'émission d'une fréquence $f=xf_H$ se fait dans les directions θ_1 et θ_2. Les énergies se propagent suivant deux directions non représentées.

Ils sont donc déterminés par l'intersection du cercle de diamètre $0W$ (W étant le point de coordonnées polaires w, β) et de la courbe Φ_x correspondante. On a donc une, ou plusieurs directions d'émission. Pour chacune d'entre elles il est possible de déterminer la direction et la vitesse du rayon.

Lorsque deux des angles θ sont confondus (ce qui se produit pour certaines valeurs de w lorsque β et f sont déterminés) *la propagation du rayon a alors lieu suivant la direction de la particule* (Fig. 7). En effet W est tel que le cercle de diamètre $0W$ est tangent à la courbe Φ_x en P. La droite inverse de ce cercle (qui est perpendiculaire à $0W$) est tangente à la

courbe C en M; la propagation a donc lieu suivant $0W$.

Dernière question que l'on peut se poser: la direction $0W$ étant toujours donnée, mais w et f ne l'étant plus, y-a-t-il des fréquences f (et des vitesses w correspondantes) pour lesquelles non seulement la propagation se ferait suivant $0W$ mais encore avec une vitesse de propagation $V=w$? La réponse est oui, mais, étant donné l'étude précédente, il faut que

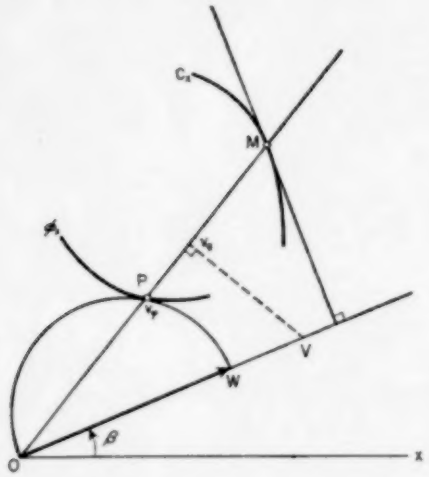


Fig. 7. Effet Čerenkov dans un milieu anisotrope. La vitesse $0W$ est telle que les deux angles θ sont confondus. Alors la fréquence $f=xf_H$ voit son énergie se propager dans la direction de la particule.

$v_g=v_\varphi$ donc $n=n'$, donc $\partial n/\partial x=0$. Or les points pour lesquels $\partial n/\partial x=0$ sont les points de contact des courbes C_x avec leur enveloppe. Les points demandés correspondent donc à une fréquence f et un angle θ tels que la courbe C_x ait avec son enveloppe un point de contact où la tangente soit perpendiculaire à $0W$. On cherche donc:

(1) Les points de l'enveloppe à tangente perpendiculaire à $0W$. Ceci donne les points M et donc les angles θ.

(2) On détermine celles des courbes C_x qui sont tangentes à leur enveloppe aux points M précédemment déterminés; ceci donne les fréquences f.

(3) L'inverse des tangentes aux points M sont des cercles dont l'intersection avec $0W$ détermine les vitesses w correspondantes (Fig. 8).

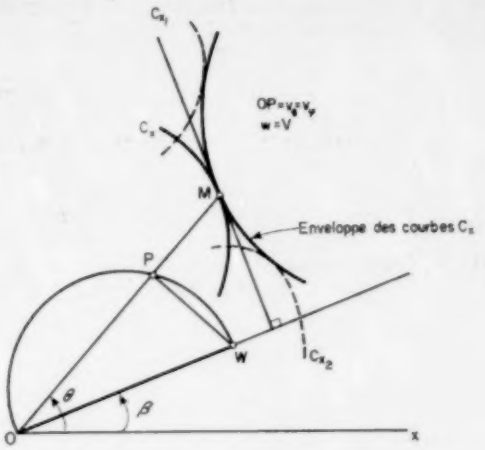


Fig. 8. Effet Cerenkov dans un milieu anisotrope. La fréquence f, pour laquelle la courbe C_x admet avec son enveloppe un point de contact où la tangente est perpendiculaire à $0W$, voit son énergie se propager suivant $0W$ et à la même vitesse que la particule.

Dans le cas des whistlers l'enveloppe des courbes C_x est une droite perpendiculaire à H_0. On n'a donc de solution pour ce problème:

(1) Que pour les particules se propageant suivant le champ magnétique.

(2) La fréquence f est alors indéterminée.

On a donc là l'explication des particularités rencontrées dans l'étude de l'effet Čerenkov des whistlers.

Il est intéressant de noter la généralité du phénomène. *Dans tout milieu anisotrope et dispersif tel que l'enveloppe des courbes d'indice soit une droite D*:

(1) *On a une propagation de toutes les fréquences le long de la droite $0x$ perpendiculaire à D à une vitesse constante.*

(2) *Des particules se propageant le long de $0x$ avec cette vitesse émettent toute leur énergie d'interaction par effet Čerenkov dans ce mode de propagation.*

BIBLIOGRAPHIE

1. L. R. O. Storey, *Phil. Trans. Roy. Soc.* **A246,** 113 (1953).
2. R. A. Helliwell, J. H. Crary, J. H. Pope et R. L. Smith, *J. Geophys. Res.* **61,** 139 (1956).
3. R. L. Smith, R. A. Helliwell et I. W. Yabroff, *J. Geophys. Res.* **65,** 815 (1960).

4. J. A. RATCLIFFE, *The Magneto-Ionic Theory and its Applications to the Ionosphere*. Cambridge University Press, Cambridge (1959).
5. G. R. A. ELLIS, *Planet. Space Sci.* **1,** 253 (1959).
6. R. M. GALLET, *Proc. Inst. Radio Engrs.* **47,** 211 (1959).
7. R. M. GALLET et R. A. HELLIWELL, *J. Res. Nat. Bur. Stand.* **D63,** 21 (1959).
8. J. R. PIERCE, *J. Appl. Phys.* **26,** 627 (1955).
9. J. V. JELLEY, *Cerenkov Radiation and its Application*. Pergamon Press, London (1958).
10. R. GENDRIN, *Comptes Rendus*, **251,** 1122 (1960).
11. R. L. SMITH, J. Research N.B.S. **64D**, 505 (1960).

Addendum

Depuis la rédaction de cet article, R. L. Smith[11] a publié une étude sur le même sujet. Sa Fig. 2, donnant l'angle maximum du rayon avec le champ magnétique, confirme le tracé de nos courbes Γ_x. D'autre part, en faisant $\cos\theta = 2\ f/f_H$ dans ses formules (7), (14) et (15) on trouve $\alpha = -\theta$ (propagation le long du champ magnétique) et $M' = 2f_0/f_H$. (M' indice de groupe du rayon, indépendant dans ce cas de la frequence). Les résultats sont donc parfaitement concordants.

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 283-285. Printed in Great Britain.

THE LUNAR MARIA

B. WARNER

University of London Observatory, Mill Hill Park, London

(*Received* 3 *February* 1961)

Abstract—The hypothesis proposed by Gold that the lunar maria are composed of dust is shown to give a detailed explanation of the present appearance of the maria. It is also shown that many bright rays may not have the extent and intensity as they had when they were formed.

1. INTRODUCTION

In 1955 T. Gold[1] put forward the proposal that the lunar maria are composed of dust, the dust having been produced by erosion of surrounding continental material. It is the purpose of this paper to discuss the development of the maria in the light of this proposal.

First let us examine the dust theory a little more closely.

Gold suggests that the dust is formed by the action of ultra-violet and X-ray radiations on the lunar surface. He estimates the rate of production of erosion materials to be of the order of 10^{-4} cm^3/cm^2/year. The dust formed would be darkened by the action of short wavelength solar radiation, whilst the continental rocks would remain bright as they are constantly being worn away and so exposing new areas.

Furthermore, Gold produces arguments to show that when formed the dust would not remain where it is but would slowly travel towards low-lying land. Thus the maria must have originally been basins towards which the dust flowed.

In the arguments that follow a critical stage in the evolution of the maria will be shown to be that when the ray craters and bright rays were formed. We therefore consider some of the properties of the ray craters.

2. DISCUSSION OF RAY CRATERS AND GENERAL HEIGHTS ON THE MOON

The ray craters were formed towards the end of crater-producing activity, and the rays are apparently light-coloured ash ejected from the parent ray craters. It is probable that the rays and ray craters are coeval.

At the time that the rays were formed the whole Moon would be covered with a thin layer of light-coloured ash; the rays themselves cover some regions to depths of probably several centimetres. All the maria were in existence when the rays were generated, as is demonstrated by the fact that many bright rays cross the maria. We can form an estimate of the time of production of the rays by reference to the recent work of Sinton[2]. From infra-red measurements on Tycho during a lunar eclipse Sinton was able to estimate the thickness of dust on the floor of Tycho. Knowing the influx rate of meteorites on the Moon he was able to put an upper limit on the age of Tycho as 10^7 years. Since the erosion of the lunar surface has been in progress since the formation of the Moon we can see that most of the present mass of eroded material was already in existence at the time the bright rays were formed.

Hence we infer that quite recently the Moon was covered with a thin layer of light-coloured ash. From that time to the present, dust will have continued to have been produced, and as it is dark it will have obscured parts of this light-coloured covering. To find where any dust produced in the last 10^7 years will have moved to we must study the general heights of the lunar surface in the neighbourhood of the maria.

The most recent and reliable contour map of the Moon is that given by Schrutka-Rechtenstamm and Hopmann[3], see Fig. 1. From their map we can extract the following information:

(a) The Mare Serenitatis is lower than the surrounding continents.

(b) The Mare Imbrium is generally lower than the surrounding continents and slopes down towards the Sinus Iridum.
(c) There is a large nearly circular area around the Sinus Iridum that is very low. This area includes part of the Mare Frigoris to the north of the Sinus Iridum.
(d) The Mare Tranquillitatis is lower than the surrounding continents, except on the western border, and there is a broad ridge running from south-east to north-west across the lower half of the Mare (roughly from the Sabine and Ritter to the Vitruvius area).
(e) There is a low area on the south-west limb, near Janssen.
(f) There is a low area on the south-east portion of the surface; the region of Tycho.
(g) The Oceanus Procellarum is lower than the surrounding continents and is lowest in its centre (to the east of Kepler).

3. THE PRESENT POSITIONS OF RECENTLY FORMED DUST

Following from observations (a) to (g) of the previous section we can make the following predictions:

(i) Dust will move into the Mare Serenitatis from the surrounding continents.
(ii) Dust will move across the Mare Imbrium towards the Sinus Iridum.
(iii) Once dust reaches the Sinus Iridum area it will remain there.
(iv) Dust will move onto the Mare Tranquillitatis, leaving an area bare across the lower half of the Mare.
(v) We expect dust to accumulate west of Janssen.
(vi) We expect dust to accumulate east of Tycho.
(vii) The Oceanus Procellarum will be gradually covered with dust, and it will be deepest to the east of Kepler.

As pointed out above, the recently formed dust will obscure the light-coloured general covering of the lunar surface. Therefore to test (i) to (vii) we examine the lunar surface with the view to finding differences in brightness in the places mentioned.

4. DISCUSSION AND INTERPRETATION OF THE PHOTOGRAPHS

As we are looking primarily for the positions of the bright ash coverings it is necessary to use Full Moon photographs. Two such photographs are to be found in the *Photographic Lunar Atlas*[4] and *The Moon*[5]. These show that there are bright regions on the maria that have been partially obscured by intruding dark material. We will examine each of the points (i) to (vii) in turn.

The centre of the Mare Serenitatis is covered with a light-coloured ash. Running around the edge of the mare is a dark border with an average width of about 80km. There is a sharp line of demarcation between the dark border and the bright interior of the Mare.

Evidently this border is the encroaching dust from the bordering continents.

The bright ray from Menelaus passes right across the Mare Serenitatis and therefore enters the dark border in two places. When it leaves Menelaus it is indistinct, but brightens sharply when it leaves the dark border. At the other (northern) side of the Mare it again becomes indistinct on reaching the dark border.

The hypothesis is, therefore, that the central regions of the Mare Serenitatis indicate the general appearance of the maria immediately after production of the bright rays.

The surface of the Mare Imbrium is generally darker than the centre of the Mare Serenitatis, and is comparable in darkness with that of the border around the latter. Exceptions are the regions south of Archimedes and Autolycus, where the ground is obviously higher. We infer that the whole of the Mare Imbrium has been covered with dust since the ray craters were formed.

The Sinus Iridum area is especially noteworthy. Here there is a large, nearly circular, area that is much darker than the rest of the Mare Imbrium. This patch corresponds exactly with the low region shown in Fig. 1 and even extends onto the Mare Frigoris to the north of the Sinus Iridum. We infer that here the dust is deeper than elsewhere and obscures nearly all of the bright ash deposits. The rays from Copernicus seem to fade out on reaching this dark area.

Fig. 1.

(*To face page* 284)

On the Mare Tranquillitatis the situation is equally interesting. The Mare gets progressively darker from east to west, with the exception of two large bright areas that lie exactly on the ridge mentioned in (d) above. These two bright areas are around Arago and to the south of Dawes; they are as bright as the centre of the Mare Serenitatis. The dust has flowed right across the Mare Tranquillitatis, being deepest on the western side and leaving the higher portions bare.

To the west of the Mare Foecunditatis, where there is a low area according to Fig. 1, are a large number of dark-floored craters, and the Maria Smythii and Marginus.

West of Janssen we also find a large number of dark-floored craters and the Mare Australe.

East of Tycho there is no accumulation of dust. The reason for this is not clear. Fig. 1 may not be reliable in this region.

The Oceanus Procellarum is mostly covered with material of the same shade as the Mare Imbrium and only in certain regions (e.g. south of Copernicus) are there brighter areas. There are indications of a darker region to the east of Kepler.

We may note that many other marial regions on the surface of the Moon have dark borders, the Maria Crisium, Nectaris, and Humorum, for example. This again may be due to dust moving down from the surrounding continents.

We have noted above that both the Menelaus and Copernicus bright rays are partially obscured by dust, and this raises the question of whether any other rays have been similarly affected.

The Tycho ray running towards Rosse seems to diminish in brightness on reaching the edge of the Mare Nectaris. Also the Tycho ray running towards Bullialdus, although starting as a broad and bright ray, disappears after travelling a short distance across the Mare Nubium. Other examples are the dimness of the Proclus rays on the borders of the Mare Crisium, and the general lack of extensive bright rays from the otherwise very bright Manilius.

A further question arising from the above discussion is why the Mare Serenitatis has not been completely covered with dust if it is not high in the centre. The reason is probably because the slopes in the Mare Serenitatis are less than those in the other maria (for instance Mare Serenitatis contains no –2000 line). The rate of movement of dust across a mare would probably depend on other factors as well, such as the quantity of dust arriving at the border of the mare.

5. CONCLUSION

The dust theory of the lunar maria, as proposed by Gold, is capable of giving a detailed and comprehensive explanation of the various shadings within the maria. The foregoing discussion shows that we are still able to see parts of the lunar surface in the state that they were left after the formation of the bright rays, but that many areas have been since modified. The rays themselves may have been changed in extent and intensity wherever they cross the maria.

Note added in proof—R. B. Baldwin has recently produced a highly detailed lunar contour map (*Sky and Telesc.* **21,** 84, 1961). All of the lunar maria are shown to be lower than the continents: the surface slopes down to the maria in most directions. The low area east of Tycho (see Fig. 1) has been replaced by an elevated area.

REFERENCES

1. T. Gold, *Mon. Not. R. Astr. Soc.* **115,** 585 (1955).
2. W. M. Sinton, *Lowell Obs. Bull.* **108** (1960).
3. G. Schrutka-Rechtenstamm and J. Hopmann, *Mathem.-naturw. Klasse, Abt II,* **167,** 283-290 (1958).
4. G. P. Kuiper, *Photographic Lunar Atlas,* Univ. of Chicago Press, Chicago (1960).
5. P. A. Moore and H. P. Wilkins, *The Moon,* Faber and Faber, London (1954).

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 286-287. Printed in Great Britain.

STRESSES IN THE MOON

GILBERT FIELDER*
University of London Observatory, London

(*Received* 25 *November* 1960)

Abstract—It is argued that a reliable theory of the stress history of the Moon should take into account several factors; and that direct observations of the Moon's surface can throw much light on this subject.

MacDonald[1] has discussed the variation of the radius of the Moon from a consideration of its thermal history. He made certain different initial assumptions about the thermal constants of the lunar rocks and the distribution and concentration of radioactivity in the Moon. These various assumptions lead to the general conclusion that the radius of the Moon increased in the Moon's early history and that this increase was followed, after a period of constancy, by a decrease of the radius at a much smaller rate. Hédervári[2] drew attention to the superficial similarity of the system of the Red Sea and African rift valleys, on the one hand, and the Ariadaeus-Hyginus and Triesnecker rilles, on the other; and he proposed[2] that the Moon is expanding as a result of the assumed decrease of the constant of gravitation with the lapse of time, and by analogy with Egyed's[3] theory of an expanding Earth. This assumption led Hédervári to propose that the surface of the Moon underwent a tensional phase.

The present writer has approached the problem from direct observation of the Moon's surface and has concluded that, whilst there may have been a tensional phase in the history of the near-surface layers of the Moon, this was not nearly as important as a compressional phase.

There is no evidence that rilles are cracks in the surface of the Moon—even though this is stated over and over again in the literature. There is evidence[4, 5, 6], on the other hand, that the Ariadaeus Rille (selected for study because of its large size and its proximity to the centre of the Moon's apparent disk) is a graben feature. Other rilles may well also be collapse features. The writer has shown[4] that, even if the Ariadaeus Rille had its origin in a fracture some distance *beneath* the surface of the Moon, this could well have been a wrench fault (and not necessarily a tensional fracture). Again, faults exposed on the surface do not necessarily indicate tension: such features as the Straight Wall may be thrust faults, for example.

Secondly, MacDonald's belief that a detailed examination of the lunar surface reveals no lateral displacements of the walls of craters cannot be sustained by the present writer. Many authors[4, 7, 8, 9, 10] have discussed the polygonal shapes of craters. Some craters, evidently the very old ones, have boundaries which deviate so much from the circular form that they are scarcely recognizable as craters (see plates in Ref. 5). How are we to explain these shapes other than as having resulted from lateral pressures?

The writer established a scale of relative time[4, 11] for the origin of the various features—again purely from direct observation of the Moon. Urey[12] contends that the largest craters are of the same order of age as the Moon itself, and this does seem to be a reasonable proposition. Certainly, the largest craters are very much older than many of the smaller craters. The majority of craters $\leqslant 20$ km in diameter which overlap, and which, therefore, were formed more recently than, the largest craters do not appear to deviate very much from the circular form.

All these observations are compatible with the assumption that there has been, in the past, a surface compressive phase in the Moon's history. Observations of the positions of points on the Moon's surface at different times seem to

* I.C.I. Research Fellow, University of London Observatory.

indicate that the Moon is not in hydrostatic equilibrium. If this is the case, stresses will be set up in the body of the Moon. An elongation of the axis of the Moon which is directed towards the Earth, and of magnitude greater than its equilibrium value, would produce a general compression of the surface.

It is important to consider all possible origins for stresses in the Moon's crust. *Uniform* stresses can arise due to:

(a) the non-equilibrium shape of the Moon
(b) tidal forces
(c) release of thermal energy by radioactive decay
(d) a possible (but not established) dependence of the "constant" of gravitation on time.

In addition, special stresses can be superimposed on the general field of stress as a result of meteoritic bombardment or igneous activity. Whilst such special stresses should not be disregarded when details are considered, the lunar grid system[13, 14] consists of families of distortions which are, in general, remarkably parallel when seen in conical projection; and this suggests that the stresses which produced those distortions were fairly uniform in direction. Also, a given family of distortions generally extends for a large distance across the Moon's surface; and this suggests that the stresses were fairly uniform in magnitude over the same distance.

MacDonald has considered case (c), and Hédervári has considered case (d). Clearly, all contributions from the factors (a) . . . . (d) to the total stress at a given time must be considered before any reliable information on the actual stress history of the Moon can be given from theoretical work alone.

REFERENCES

1. G. J. F. MacDonald, *Planet. Space Sci.* **2**, 249 (1960).
2. P. Hédervári, *Die Sterne*. **36**, (1960); *Hung. Phys. Rev.*, **8**, 261 (1960).
3. L. Egyed, *Acta Geol., Acad. Sci., Hung.*, **4**, 43 (1956).
4. G. Fielder, *J. Int. Lunar Soc.*, **1**, 166 (1960).
5. G. Fielder, *Sky and Telesc.* **19**, 334 (1960). (This paper is an abridged version of Ref. 4).
6. G. P. Kuiper, in *Vistas in Astronautics*, Vol. II, part V. (Ed's. M. Alperin and H. F. Gregory), Pergamon, London (1959).
7. P. Puiseux, *La terre et la lune* Gauthier-Villars Paris (1908).
8. F. C. Lamech, *Carte général de la lune*, Girard, Barrère et Thomas, Paris (1946).
9. K. Graff, in *Handb. Astrophys.* **4**, p. 406 (1929).
10. D. Alter, *Publ. Astr. Soc. Pacific*, **68**, 437 (1956).
11. G. Fielder, *Publ. Astr. Soc. Pacific*, **70**, 308 (1958).
12. H. C. Urey, *Endeavour*, **19**, 87 (1960).
13. G. Fielder, *J. Brit. Astr. Ass.* **67**, 314 (1957).
14. G. Fielder, *J. Brit. Astr. Ass.* **67**, 60 (1957).

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 288-291. Printed in Great Britain.

EXCITATION OF PLASMA WAVES BY BODIES MOVING IN IONIZED ATMOSPHERE*

K. P. CHOPRA

Polytechnic Institute of Brooklyn, Aerodynamics Laboratory, Freeport, N.Y.

(*Received* 23 *December* 1960)

Abstract—A body moving in an ionized atmosphere acquires an electric charge through the processes of accretion of charged particles and emission of electrons by high energy photons. The moving charged body may then interact with the charged particles of the atmosphere and any pervading magnetic field to excite plasma waves. Of particular interest is the situation in which the body collects an ionized cloud in front of it. The motion of this ionized cloud in the atmosphere induces an electrostatic instability and causes a column of ionized gas to move ahead of the body. The electrostatic instability is conducive to the excitation of electrostatic oscillations which if already present are further enhanced. A magnetic field along the direction of motion assists in the formation of the ionized cloud. If the pervading magnetic field is of suitable weak strength, it may excite extraordinary electromagnetic waves. A pervading transverse magnetic field of suitable strength may cause the excitation of magnetohydrodynamic waves.

1. FORMATION OF AN ION CLOUD

Let us consider the rapid motion of a body in a rare ionized atmosphere pervaded by a weak magnetic field. We assume that

(i) the atmosphere is composed of electrically neutral particles, ions and electrons which are in local thermodynamic equilibrium corresponding to the gas-kinetic temperature T,

(ii) the characteristic linear size, say radius a, of the test body is small compared to the gas-kinetic mean free path λ, and

(iii) the speed v of the test body is large compared to the mean thermal speed v_i of ions and neutral particles, and is small compared to the mean thermal speed v_e of electrons.

We will adopt $v \sim 8 \times 10^5$ cm/sec as the representative speed of the test body, and $m_i \sim 14\, m_p$ (with m_p as the proton mass) as the representative mass of ions and neutral particles. The neutral particles and ions are incident to the test body with the speed of the test body, the impact energy being of the order of 5 electron volts.

About one half of the incident particles are reflected back with the impact speed with respect to the body. The other half of the impinging particles are trapped in the metal surface. Their kinetic energy relative to the body is given up to the body as heat. The surface temperature is thereby raised. This promotes the vaporization of the test body. Because of the low impact energies involved, the phenomenon of sputtering does not play any role in the particle-surface interactions.

The ratio of mass of vapors to that of incident particles is of the order of 3. The vapor molecules of mass μ leave the surface with a speed of the order of $(kT_s/\mu)^{1/2}$ with respect to the test body. Here k is the Boltzmann constant and T_s is the surface temperature (absolute scale) of the test body. In the coordinate system at rest with the atmosphere, therefore, a vapor possesses kinetic energy of 10 eV or between 20–30 eV according as the test body is made of *Al* or *Fe*. On the other hand, the reflected particles have energies of the order of 10 eV in this system.

In collisions with atmospheric particles, or between themselves, the re-emitted (reflected

* Supported in part by the United States Air Force Office of Scientific Research Contract No. AF-49(638)-445, and presented at the Second Annual Meeting of the Plasma Physics Division of the American Physical Society held at Gatlinburg, Tenn. on 2–5 November 1960.

or vapor) molecules may produce dissociation, excitation and ionization. At kinetic energies involved in the present case, the atmospheric particles are subject only to dissociation. Their proportionately slight ionization caused in collisions with the re-emitted particles may be attributed to the process of charge exchange.

Again, through the accretion of charged particles and photo-electric effect, the test body acquires an electric charge[1]. The atmospheric ions and electrons, under the influence of Coulomb force, describe hyperbolic orbits about the test body. The particles scattered in this manner possess kinetic energy of the order of 10 eV.

The role of collisions differs substantially at different distances from the surface of the body. Henceforth we will refer to the reflected, scattered, vapor or secondary particles as re-emitted particles. In collisions very close to the test body, either one or both of the impinging and re-emitted particles hit the surface. On the other hand, collisions at great distances shield the body from impacts with the oncoming atmospheric particles. Furthermore, at each collision ahead of the test body, the re-emitted particle imparts a certain velocity to the atmospheric particle in the forward direction. This continues until the forward motion of the re-emitted particle is destroyed and it is turned back or sideways. A series of collisions of this type build up an ionized cloud or avalanche of particles moving ahead of the test body. The cloud expands laterally through diffusion processes. A longitudinal magnetic field pervading in the atmosphere has a tendency to arrest the lateral diffusion of charged particles, and to constrain them in the forward moving column. It also assists in trapping scattered particles in the cloud.

2. COUNTER-STREAMING

The penetration of the ionized cloud in the atmosphere may be treated as the interpenetration of two ionized streams. For the purposes of the following sections, we will assume that the test body is moving in an atmosphere composed only of charged particles, and acquires an ionized cloud in its front. We will also assume that all particles are singly charged. At the particle impact energies involved here it is also reasonable to assume that the densities of the frontal cloud and the atmosphere are equal. Then the motion of the ionized cloud in the atmosphere is analogous to the interpenetration of two fully ionized streams of equal density penetrating each other with a relative velocity equal to the speed of the test body.

Following the analysis of the interpenetration of two ionized streams[2], the dispersion relation governing the sinusoidal disturbances may be written as

$$\frac{\omega_p^2}{(\omega-\frac{1}{2}kv)^2}+\frac{\omega_p^2}{(\omega+\frac{1}{2}kv)^2}=1 \tag{1}$$

where $\omega_p=(n_e e^2/\pi m_e)^{1/2}$ is the plasma frequency and $k\sim 1/\lambda$ is the wave number. This relation reduces to a biquadratic in ω,

$$\omega^4-2(\omega_p^2+\tfrac{1}{4}v^2k^2)\omega^2+\tfrac{1}{4}v^2k^2(\tfrac{1}{4}v^2k^2-2\omega_p^2)=0 \tag{2}$$

For given values of k, v and ω_p, it follows that (2) yields real values for ω^2, and for $k^2v^2-8\omega_p^2<0$, one value of ω^2 is negative. This value of ω^2 gives a pair of complex roots of ω and results in the instability of the medium. Piddington[3] has discussed a simple mechanism wherein this instability leads to the growth of the ionized cloud. An ionized cloud, however low its initial density, will, if its linear dimension is large enough, continuously increase in density with time. The criterion for growth is that the initial linear dimension of the cloud must exceed $\lambda_{\min}$ given by

$$\lambda_{\min}\sim\frac{1}{k_{\max}}\sim\left(\frac{\pi mv^2}{8n_e e^2}\right)^{1/2} \tag{3}$$

The cloud sustains itself through the balance of mechanisms of its growth and dissipation.

3. ELECTROMAGNETIC WAVES

Let us now assume that a weak magnetic field H pervades the atmosphere. The charged particles of the atmosphere gyrate around the magnetic lines of force with cyclotron frequency $\omega_e=eH/2\pi mc$. Because of their lighter mass, the electrons spiral more freely than the ions. Let us now consider that the test body, with its

frontal cloud, is moving parallel to the magnetic field with speed v. The charged particles in the ionized cloud interact with the gyrating charged particles to excite slow electromagnetic waves. The electric vector of the waves rotates in the same sense as the electrons gyrate. These waves propagate with a phase velocity slightly less than the speed of the test body, and in this manner the phenomenon bears a resemblance to Čerenkov radiation effect[3]. The phenomenon is also analogous to the principle on which the traveling wave tubes[4] operate, the charged particles of the cloud providing their kinetic energy for the electromagnetic waves. Chopra and Singer[1] described this mechanism, with the participation of only the test body, as one of the several mechanisms contributing to the drag of the test body. Gallet[5] has invoked this mechanism, involving very fast single test particles carrying an electric charge, to explain very low frequency emissions in the atmosphere of the earth. It may be remarked that the frontal cloud is more efficient here because its particles bunch together to enhance the radiation effect.

The dispersion relation governing these waves is given by

$$\left(\frac{v_{ph}}{c}\right)^2 = \frac{\omega\,(\omega-\omega_{ce})}{\omega^2-\omega\,\omega_{ce}-\omega_p{}^2} \tag{4}$$

which, on setting $v_{ph}=v$, reduces to the quadratic in ω,

$$\omega^2-\omega\,\omega_{ce}+\omega_p{}^2\frac{v^2}{c^2-v^2}=0. \tag{5}$$

The solution of the last equation is

$$\omega=\tfrac{1}{2}\omega_{ce}\left[1\pm\left\{1-\frac{4m_e}{m_i}\left(\frac{v}{v_A}\right)^2\right\}^{1/2}\right], \tag{6}$$

where $v_A{}^2=H^2/4\pi n_i m_i$ is the square of the Alfvén speed. Let us write $\alpha=4\,(m_e/m_i)\,(v/v_A)^2$. For given values of ω_{ce} and α, there are in general two possible values of the wave frequency. For $\alpha\ll 1$, one root corresponds approximately to the electron cyclotron frequency and the other root corresponds to a very very low frequency equal to $\alpha/2$. For $\alpha=1$, the two roots are each equal to half the value of the electron cyclotron frequency. For $\alpha>1$, the wave propagation is not possible. Gyrating ions can also take part in the interaction. A similar set of wave frequencies depending on the ion cyclotron frequency ω_{ci} and $\alpha'=(2v/v_A)^2$ are possible. When $v<\frac{1}{2}v_A$, the ions produce lower frequencies than those produced by electrons. In the presence of neutral particles, α and α' are replaced by $(n_e/n)\alpha$ and $(n_e/n)\alpha'$ respectively where n_e/n is the degree of ionization.

4. MAGNETOHYDRODYNAMIC WAVES

In this section we will consider the case of a test body and its frontal cloud moving in a direction transverse to the pervading magnetic field. This transverse motion of the ionized cloud interacts with the pervading magnetic field to produce magnetohydradynamic waves. Ferraro[6] has demonstrated that the propagation of magnetohydrodynamic waves is in fact possible in an ionized cloud of very low density. He obtained the dispersion relation

$$(\omega-\omega_{ce})\,(\omega+\omega_{ci})\,(\omega^2-k^2c^2)=\omega^2\omega_{pv}{}^2 \tag{7}$$

where $\omega_{pv}=(n_e e^2/\pi\nu)^{1/2}$ is the plasma frequency corresponding to the reduced mass $\nu=m_e m_i/(m_e+m_i)$. Subject to the condition,

$$\omega^2<\omega_{ce}\omega_{ci}<\omega_{pv}{}^2 \tag{8}$$

these waves propagate with a phase velocity equal to the Alfvén speed $v_A=H_o/[4\pi n_e\,(m_e+m_i)]^{1/2}$. For $\omega<\omega_{ce}$, these waves propagate without dissipation as if the ionized atmosphere possessed infinite electrical conductivity. The condition (8) combined with

$$H^2<4\pi n_e\,(m_e+m_i)c^2 \tag{9}$$

determines that the waves are of the magnetohydrodynamic type. The last equation states that the Alfvén speed must be less than the velocity of light. It also translates the condition[7] that an ionized gas is able to move freely through a transverse magnetic field. It sets an upper limit to an initial magnetic field necessary for the excitation of these waves.

Note added in proof—Since this paper was prepared, a paper by Lighthill[8] has appeared in which he asserts that a space vehicle traversing the ionosphere cannot excite magnetohydrodynamic waves because

the Alfvén speed far exceeds the speed of the space vehicles; but it can only excite the slower sound waves. It has recently been demonstrated by Howard Chang[9] that Lighthill's assertion is correct when applied to sound waves, and is false when applied to the magnetohydrodynamic waves. It may be mentioned here that the criterion used by Lighthill does not apply to wave phenomena in general, and the difference between propagation and excitation of waves may not be overlooked. Whereas the propagation of waves is a characteristic property of the medium, the excitation of waves depends very definitely on a specific excitation mechanism. A particle scattered by the space vehicle in the forward direction will in general suffer several successive Coulomb collisions, because, after suffering each Coulomb collision, it is trapped by a magnetic line of force a suitable distance ahead of it and describes a cyclotron gyration which brings it once again closer to the approaching space vehicle only to suffer another Coulomb interaction. This process continues until the particle either hits the vehicle or is shoved sideways and gets completely out of the way of the space vehicle. At each successive Coulomb scattering, the particle gains momentum which results in a jump of its guiding center from one line of force to another. This jump of the guiding center is somewhat equivalent to the plucking of a stretched elastic string, and hence causes the excitation of the Alfvén waves.

REFERENCES

1. K. P. Chopra and S. F. Singer, *Drag of a Sphere Moving in a Conducting Fluid*, University of Maryland Physics Department Technical Report 97 (January 1958); pre-printed, 1958 *Heat Transfer and Fluid Mechanics Institute*, p. 166 (1958).
2. J. H. Piddington, *Phil. Mag.* **3**, 1241 (1958).
3. P. A. Cerenkov, *Phys. Rev.* **52**, 378 (1937).
4. J. A. Pierce, *Traveling Wave Tube*, Van Nostrand, New York, N.Y. (1950).
5. R. M. Gallet, *Proc. Inst. Radio Engrs.* **47**, 211 (1959).
6. V. C. A. Ferraro, *Proc. Roy. Soc.* **A233**, 310 (1955).
7. F. Hoyle, *Recent Researches in Solar Physics*, p. 16. Cambridge University Press, Cambridge (1949).
8. M. J. Lighthill, *J. Fluid Mech.*, **9**, 465 (1960).
9. H. H. C. Chang, (Private Communication).

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 292-298. Printed in Great Britain.

ELECTRO-HYDROMAGNETIC WAVES IN A FULLY IONIZED GAS - II

K. D. COLE

Antarctic Division, Department of External Affairs, Melbourne

(*Received* 28 *December* 1960)

Abstract—The propagation of hydromagnetic and low frequency radio waves in all directions in a fully ionized gas containing a magnetic field is examined. For longitudinal and transverse propagation the addition of one extra term in the magneto-ionic formulae (without collisions) accounts for the presence of heavy ions.

The partition of energy of disturbance between kinetic (K) and magnetic (M) for longitudinal propagation of all frequencies is given by

$$K = V^2 \frac{k^2}{\omega^2} M \left(1 - \frac{\omega^2}{k^2 c^2}\right)^2$$

where V is the Alfvén speed. Thus approximate equipartition may exist for some audio- and radio-frequencies in the Earth's exosphere.

Some errors in Paper I of this series are corrected.

1. INTRODUCTION

An analysis is reported describing propagation in all directions of waves of hydromagnetic and radio frequencies in a fully ionized gas. Well-known magneto-ionic formulae are deduced as a special case. Alfvén's theorem for the partition of energy of disturbance is generalized to all frequencies.

In a previous paper[1] (*E.H.M. Waves in a Fully Ionized Gas—I*, hereafter called *Paper I*), dispersion relationships (Equations (15) and (16) of that paper) were derived purporting to describe the propagation of waves of all frequencies orthogonal to a magnetic field in a fully ionized gas. The author now considers these dispersion relationships to be inadequate. Equation (9) of *Paper I*, one of the equations from which they were derived, tacitly assumes that the divergence of conduction current is zero. Such an assumption is not valid for all frequencies (c.f. reference [2]). In the present paper this restriction is removed and some new results deduced.

2. THE DISPERSION EQUATION

In this paper the c.g.s. system of units will be used throughout. In Maxwell's equations the dielectric constant and magnetic permeability are each taken as unity. Thus

$$\text{curl}\,\mathbf{H} = \frac{4\pi}{c}\mathbf{j} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} \tag{1}$$

$$\text{curl}\,\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} \tag{2}$$

In an ionized gas the conduction current density j due to a weak oscillating applied electric force F of frequency ω in the presence of a magnetic field H is

$$\mathbf{j} = s_0 \mathbf{F}_{//} + s_1 \mathbf{F}_{\perp} + s_2 \frac{\mathbf{H} \times \mathbf{F}_{\perp}}{H} \tag{3}$$

where $\mathbf{F}_{//}$, $\mathbf{F}_{\perp}$ are components of $\mathbf{F}$ parallel and orthogonal to $\mathbf{H}$. It was shown in *Paper I* that in a fully ionized gas

$$\begin{aligned} s_0 &= \frac{ne^2\tau}{m} \cdot \frac{1}{1 + i\omega\tau} \\ s_1 &= \frac{ne^2\tau}{m} \cdot \frac{1 + i\omega\tau}{(1 + i\omega\tau)^2 + p^2\tau^2} \\ s_2 &= \frac{ne^2\tau}{m} \cdot \frac{p\tau}{(1 + i\omega\tau)^2 + p^2\tau^2} \end{aligned} \tag{4}$$

where

n = electron density,

e, m = charge, mass of an electron,

τ = mean time between collisions of an electron with ions,
p = gyrofrequency of electrons,
$i = \sqrt{-1}$.

For present considerations $\mathbf{F}$ has two parts: (*i*) an electric field $\mathbf{E}$ as observed in a frame of reference at rest with respect to the centre of of mass of an element of gas, and (*ii*) a dynamo field $\mathbf{v} \times \mathbf{H}/c$ due to the movement of the centre of mass of an element of gas across the magnetic field with velocity $\mathbf{v}$. Thus

$$\mathbf{F} = \mathbf{E} + \mathbf{v} \times \mathbf{H}/c \tag{5}$$

For weak disturbances the hydrodynamic equations yield[2]

$$\frac{\partial^2 \mathbf{v}}{\partial t^2} - U^2 \operatorname{grad} \operatorname{div} \mathbf{v} + \frac{V^2}{H_0^2} \frac{\partial}{\partial t} (\mathbf{H}_0 \times \operatorname{curl} \mathbf{H}) = 0 \tag{6}$$

where $V^2 = H_0^2/4\pi\rho$, ρ being the mass density of the gas. V is the Alfvén speed. In equation (6) curl $\mathbf{H}$ is an approximation for $4\pi\mathbf{j}/c$. Thus (by equation (1)) results obtained by it are valid only so long as phase speeds derived from it are very much less than the speed of light. This is generally so for hydromagnetic waves and a large class of audio- and radio-frequency waves, as will be seen in what follows. It will be assumed in this paper that the second term in equation (6) is negligible. This means that either the gas is assumed incompressible ($\operatorname{div} \mathbf{v} = 0$) or that U (the speed of sound in the gas) is small compared to other phase speeds in the medium.

Using the above equations and restrictions we seek solutions for weak disturbances of the form $\exp i(\omega t - k\xi)$ propagated in the fixed direction 0ξ (in the $x0z$ plane) of a rectangular cartesian system of coordinates $0(x, y, z)$. The resident magnetic field ($\mathbf{H}_0$) is taken to be in the z direction. The normal (0ξ) to the wave-front makes an angle ψ with the magnetic field $\mathbf{H}_0$. The equations (1), (2) and (6) give nine linear equations in v_x, v_y, v_z, E_x, E_y, E_z, H_x, H_y, H_z. These variables may be eliminated to yield the following dispersion relationship valid for frequencies such that $\omega/k \ll c$.

$$Ak^4 + Bk^2 + C = 0 \tag{7}$$

where

$$A = \left(4\pi s_0 i\omega - \frac{\omega^2}{c^2}\right)\cos^2\psi \left\{\left(1 - \frac{V^2 4\pi s_1 i}{\omega}\right)^2 + \left(\frac{V^2 4\pi s_2 i}{\omega}\right)^2\right\} + \sin^2\psi \left\{\left(4\pi s_1 i\omega - \frac{\omega^2}{c^2}\right)\left(1 - \frac{V^2 4\pi s_1 i}{\omega}\right) + V^2 (4\pi s_2)^2\right\}$$

$$B = \left(4\pi s_0 i\omega - \frac{\omega^2}{c^2}\right)(1 + \cos^2\psi) \left\{\left(4\pi s_1 i\omega - \frac{\omega^2}{c^2}\right)\left(1 - \frac{V^2 4\pi s_1 i}{\omega}\right) + V^2 (4\pi s_2)^2\right\} + \sin^2\psi \left\{\left(4\pi s_1 i\omega - \frac{\omega^2}{c^2}\right)^2 + (4\pi s_2 i\omega)^2\right\}$$

$$C = \left(4\pi s_0 i\omega - \frac{\omega^2}{c^2}\right)\left\{\left(4\pi s_1 i\omega - \frac{\omega^2}{c^2}\right)^2 + (4\pi s_2 i\omega)^2\right\}$$

This general equation may be used for any gas satisfying the hydrodynamic conditions specified above. For a lightly ionized dense gas the conductivity coefficients s_0, s_1, s_2 would each be the sum of coefficients representing conduction contributions from each type of charged particle, and ρ the total mass density.

Of particular interest is the transmission of low frequency electro-hydromagnetic waves in the gas above the earth's ionosphere. The propagation of whistlers is likely to involve hydromagnetic as well as electromagnetic effects at some frequencies. The simpler magneto-ionic formulae do not take hydromagnetic effects to account.

Let us consider a rarefied fully ionized gas ($\tau \to \infty$) such as may exist in the earth's exosphere and interplanetary space. As $\tau \to \infty$,

$$\begin{aligned} 4\pi s_0 i\omega &\to \frac{\omega_p^2}{c^2} = \frac{\omega^2}{c^2} x \\ 4\pi s_1 i\omega &\to \frac{\omega_p^2}{c^2}\frac{\omega^2}{\omega^2 - p^2} = \frac{\omega^2}{c^2}\frac{x}{1 - y^2} \\ 4\pi s_2 &\to \frac{\omega_p^2}{c^2}\frac{p}{p^2 - \omega^2} = -\frac{\omega}{c^2}\frac{xy}{1 - y^2} \end{aligned} \tag{8}$$

where $\omega_p^2 = 4\pi n e^2/m$ and x and y are defined $x = \omega_p^2/\omega^2$, $y = p/\omega$. Let $V/c = \beta$. (In passing it is noted that in a fully ionized gas $\beta^2\omega_p^2 = \frac{m}{M} p^2$ where M is the ionic mass.)

Equation (7) becomes for this case $\left(\mu^2=\frac{c^2k^2}{\omega^2}\right)$

$$A'\mu^4+B'\mu^2+C'=0 \tag{9}$$

where

$$A'=(x-1)\left(1-\frac{\beta^2x}{1+y}\right)\left(1-\frac{\beta^2x}{1-y}\right)\cos^2\psi + \left(\frac{x+\beta^2x-\beta^2x^2}{1-y^2}-1\right)\sin^2\psi$$

$$B'=(x-1)\left(\frac{x+\beta^2x-\beta^2x^2}{1-y^2}-1\right)(1+\cos^2\psi) + \left(1-\frac{x}{1+y}\right)\left(1-\frac{x}{1-y}\right)\sin^2\psi$$

$$C'=(x-1)\left(1-\frac{x}{1+y}\right)\left(1-\frac{x}{1-y}\right).$$

The corresponding well-known magneto-ionic formula is deducible from this by placing $\beta=0$ corresponding to ions of infinite mass. Equation (9) now replaces equations (15) and (16) of *Paper I*.

When $\psi=0$, we find from (9)

$$\mu^2=\frac{1\pm y-x}{1\pm y-\beta^2x} \tag{10}$$

and when $\psi=\pi/2$ we find

either $\mu_E^2=1-x$ (11a)

or $\mu_0^2=\dfrac{x(x-1)-(x+y^2-1)}{\beta^2x(x-1)-(x+y^2-1)}$ (11b)

Equations (10), (11a) and (11b) should be compared with equations (24), (25) and (26) in Chapter VI of Mitra's book[3]. The effect of heavy ions is seen to be catered for by the inclusion of a single term (involving β) in the simpler magneto-ionic formulae.

The hydromagnetic waves may be described by equation (9) by allowing $\omega\to0$. In this case equation (9) becomes

$$(\beta^2\cos^2\psi\cdot\mu^2-1)(\beta^2\mu^2-1)=0 \tag{12}$$

This yields the phase velocities of $\beta c\cos\psi$ for the extraordinary hydromagnetic wave and βc for the ordinary hydromagnetic wave.

For purposes of computation of μ^2 from equation (9) the quantities A', B', C', may be written as functions of ω_p/ω, m/M, ψ, and $R=p/\omega_p$. For a proton-electron gas values of μ^2 have been computed for a range of values of ω_p/ω for $\psi=0°$, 15°, 30°, 45°, 60°, 75°, 90° and for R values of 0·1, 1·0, and 10·0.

The variation of dispersion characteristics with ψ for $R=0{\cdot}1$ is illustrated in Figs. 1–4, and with R for $\psi=0$ in Figs. 1, 5, 6. The graphs (not reproduced here) for $R=1{\cdot}0$, 10·0 (and $\psi=15°$, 45°, 90°) are qualitatively similar to those for $R=0{\cdot}1$. For a wave in general $k=a+ib$ where a and b are real. Thus $\mu^2\,(=c^2k^2/\omega^2)$ positive (full lines in the graph) characterizes a progressive wave of phase velocity $\omega/a\,(=c/\mu)$ and zero attenuation (viz. $b=0$). μ^2 negative (dashed lines in the figures) characterizes an evanescent wave with attenuation distance $1/b\,(=c/i\mu)$. The waves are transmitted in two frequency blocks—high frequency ($\omega>\omega_p$ electromagnetic not graphed) and low frequency (hydromagnetic). There is a block of frequencies between the two which does not propagate and suffers attenuation. Fig. 7 shows a plot of the upper

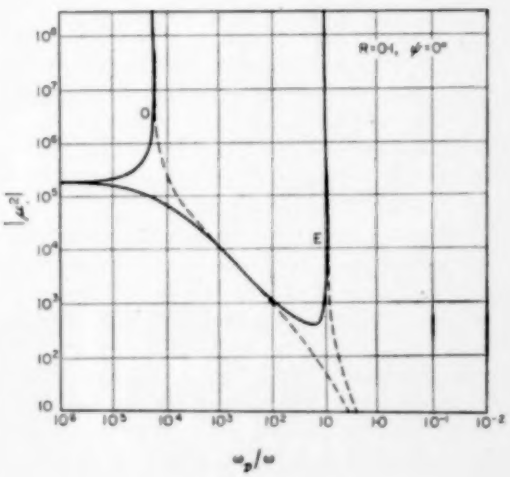


Fig. 1. Plot of $|\mu^2|=\left|\frac{c^2k^2}{\omega^2}\right|$ against ω_p/ω for a fully ionized gas. Full line denotes positive values of μ^2; broken line, negative values of μ^2. R= ratio of gyro to plasma frequency. ψ=angle between wave normal and magnetic field.

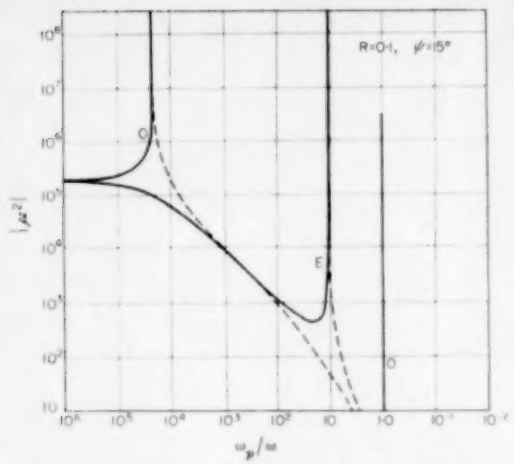


Fig. 2. Plot of $|\mu^2|=\left|\frac{c^2k^2}{\omega^2}\right|$ against ω_p/ω for a fully ionized gas. Full line denotes positive values of μ^2; broken line, negative values of μ^2. R = ratio of gyro to plasma frequency. ψ = angle between wave normal and magnetic field.

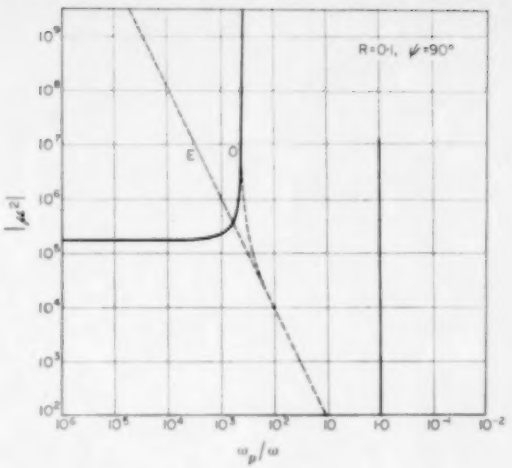


Fig. 4. Plot of $|\mu^2|=\left|\frac{c^2k^2}{\omega^2}\right|$ against ω_p/ω for a fully ionized gas. Full line denotes positive value of μ^2; broken line, negative values of μ^2. R = ratio of gyro to plasma frequency. ψ = angle between wave normal and magnetic field.

and lower limits of the missing band of frequencies as a function of ψ for the case $R=0{\cdot}1$. There should be such a missing band of frequencies from extraterrestrial sources depending on the properties of the interplanetary medium. The band is narrowest at $\psi=0$. The lower (ω_L) end frequencies of the band for a given ψ and a given wave may be found from equation (9) by putting $C/B=0$. It follows that when $\psi=0$,

$$\omega_{L,\psi=0}=\sqrt{p^2/4+\beta^2\omega_p^2}\pm p/2.$$

Now $\beta\omega_p=p\sqrt{m/M}$. Thus for a proton-electron gas $\beta\omega_p\ll p/2$. In interplanetary space it is likely that $\omega_p\gg p$. Hence for the interplanetary medium $\omega_{L,\psi=0}\approx p$. The highest frequency of the missing band would be approximately ω_p. If the magnetic field of interplanetary space is 10^{-4} gauss[7] and the electron density of order 1 per cm^3 the missing band of frequencies would be from 200 c/s to 10 kc/s.

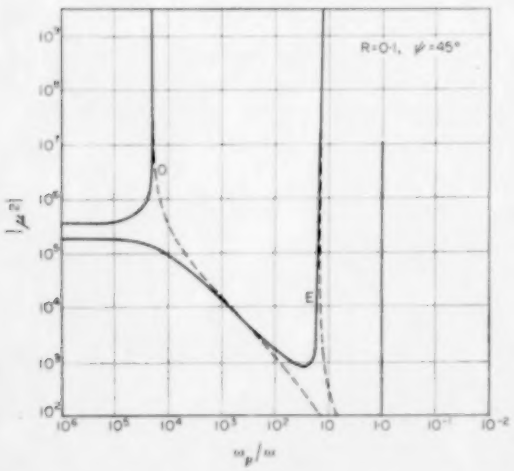


Fig. 3. Plot of $|\mu^2|=\left|\frac{c^2k^2}{\omega^2}\right|$ against ω_p/ω for a fully ionized gas. Full line denotes positive values of μ^2; broken line, negative values of μ^2. R = ratio of gyro to plasma frequency. ψ = angle between wave normal and magnetic field.

It is to be noticed in passing that the hydromagnetic wave which propagates with the same phase velocity in all directions is the counterpart at lower frequencies of the "ordinary" electromagnetic wave. The hydromagnetic wave which travels with the phase velocity of $V\cos\psi$ is the counterpart of the "extraordinary" electromagnetic wave. Astrom's[4]

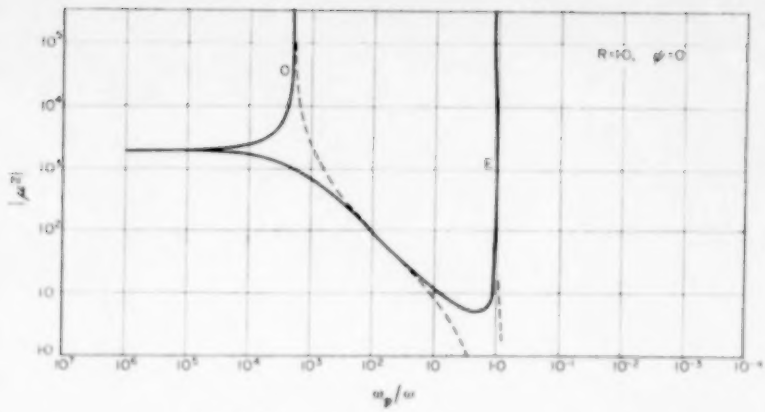


Fig. 5. Plot of $|\mu^2| = \dfrac{c^2 k^2}{\omega^2}$ against ω_p/ω for a fully ionized gas. Full line denotes positive values of μ^2; broken line, negative values of μ^2. R = ratio of gyro to plasma frequency. ψ = angle between wave normal and magnetic field.

use of the terms "ordinary" and "extraordinary" hydromagnetic waves is consistent with this. However, these labels for hydromagnetic waves are interchanged in Piddington's[5] work which the present writer has hitherto followed[1, 6] in reference to these waves. No error in physical interpretation has been made, but merely an inconsistency of labelling. For consistency authors should in the future follow Astrom's use of the terms.

2.1. *Polarization*

Let us denote the component of perturbation magnetic field in the XOZ plane by h_p. The magnetic polarization is indicated by the ratio $r = h_y/h_p$. It follows from equations (1) to (6) that

$$r = i \frac{\dfrac{xy}{1-y^2}(1-\beta^2\mu^2)\cos\psi}{\mu^2\left(1-\dfrac{\beta^2 x}{1-y^2}\right)\cos^2\psi - \left(1-\dfrac{x}{1-y^2}\right)\left(1-\dfrac{\mu^2\sin^2\psi}{1-x}\right)} \tag{14}$$

Thus in general the magnetic polarization is elliptical. In the special case of $\psi = 0$ (longitudinal propagation) we may show with the aid

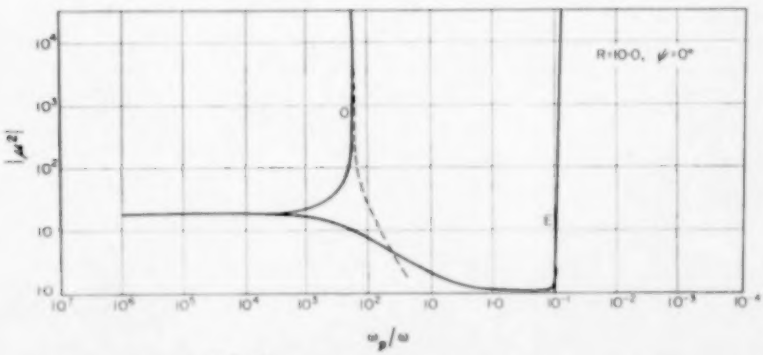


Fig. 6. Plot of $|\mu^2| = \dfrac{c^2 k^2}{\omega^2}$ against ω_p/ω for a fully ionized gas. Full line denotes positive values of μ^2; broken line, negative values of μ^2. R = ratio of gyro to plasma frequency. ψ = angle between wave normal and magnetic field.

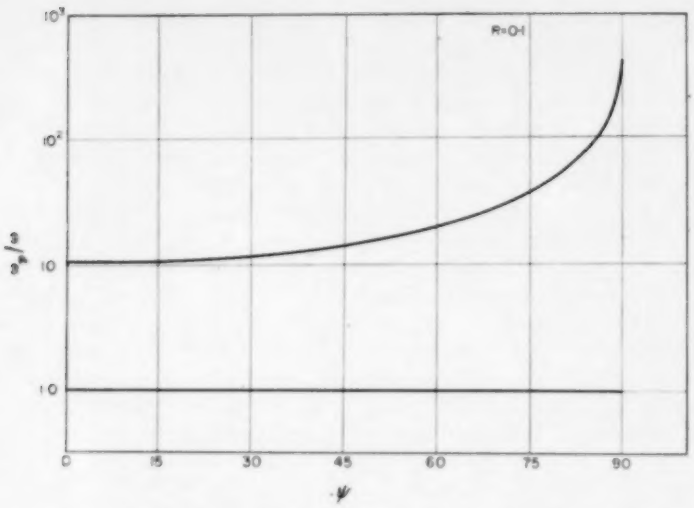


Fig. 7. Plot of ω_p/ω against ψ for the upper and lower limits of the missing band of frequencies in a fully ionized gas. R = ratio of gyro to plasma frequency.

of equation (10) that $r/i = \pm 1$, where $i = \sqrt{-1}$. For $\psi = \pi/2$ (transverse propagation) it may be shown using equations (1) to (6) that $h_x = h_z = 0$ for the extraordinary wave (c.f. equation 11a), and that $h_x = h_y = 0$ for the ordinary wave (c.f. equation 11b).

2.2. *Partition of energy between electromagnetic and hydrodynamic field*

The partition of energy of disturbance between electromagnetic and hydrodynamic field may be discussed throughout the whole range of frequencies ($\omega = 0$ to ∞) with the aid of the following equation which is appropriate to an incompressible gas or to a gas in which the speed of sound is low.

$$\rho \frac{\partial \mathbf{v}}{\partial t} = \frac{\mathbf{j} \times \mathbf{H}}{c} = \frac{1}{4\pi}\left(\operatorname{curl} \mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}\right) \quad (15)$$

Let $R(A)$ denote the real part of a complex quantity A and A^* its complex conjugate. The time average kinetic energy of disturbance is given by

$$K = \frac{\rho}{2}[\{R(v_x)\}^2 + \{R(v_y)\}^2]_{\text{av.}} = \frac{\rho}{4} R[v_x v_x^* + v_y v_y^*]$$

From equations (2) and (15) it follows that when k is real

$$K = \frac{V^2}{8\pi}\frac{k^2}{\omega^2}[\{R(h_p)\}^2 + \{R(h_y)\}^2 \cos^2\psi]_{\text{av.}} \left(1 - \frac{\omega^2}{k^2c^2}\right)^2 \quad (16)$$

where $h_p = h_z/\sin\psi$.

Thus when $\psi = 0$,

$$K = V^2 \frac{k^2}{\omega^2} M \left(1 - \frac{\omega^2}{k^2c^2}\right)^2 \quad (17)$$

where M is the time average magnetic energy of disturbance. Moreover, it is seen from equation (16) that K/M is of the same order for all values of ψ.

Equation (16) is a generalization to all frequencies of Alfvén's partition theorem[8]. It replaces equation (20) of *Paper I* which is in error by the factor $(1 - \omega^2/k^2c^2)^2$. Approximate equipartition of energy is seen to exist when $\omega/k \approx V \ll c$. *This condition exists not only for hydromagnetic waves but also for some low frequency radio waves.* This may be seen by inspection of Figs. 1, 2, 3, 5, 6. This suggests a common mode of generation may be appli-

cable to hydromagnetic waves and such radio frequency waves, viz. mechanical vibration of the gas. Thus the mechanical impact of hydromagnetic waves on an inhomogeneity of scale less than their wavelength may cause the generation of radio frequency waves through non-linear mechanical response. This would explain in a natural way the increase of dawn chorus and hiss with increased geomagnetic disturbance which (in *Paper I*) is attributed to the propagation of hydromagnetic waves all the way from the sun. Inspection of Figs. 1–6 shows that audio frequency radio waves are in the hydromagnetic block of frequencies. In passing it is noted that the nose-whistler characteristic appears in Figs. 1, 2, 3, 5, 6.

3. DISCUSSION

Equation (17) invalidates one of the assertions of *Paper I* that there is kinetic energy of disturbance associated with waves such that $\omega/k=c$. Nevertheless, an important point is clarified that when $\omega/k \gtrless c$ there is kinetic as well as magnetic energy of disturbance. The kinetic energy is not insignificant for some frequencies other than hydromagnetic ones, and may have important consequences to theories of radio-noise generation in plasmas.

It is to be noted that three frequencies emerge as important for the propagation of e.h.m. waves across a fully ionized binary gas. The plasma frequency $\omega_p=\sqrt{4\pi ne^2/m}$, the gyrofrequency of electrons $p=He/mc$, and the new frequency $\beta\omega_p$. Neither the gyrofrequency of ions (He/Mc) nor the "ion" plasma frequency $\sqrt{4\pi ne^2/M}$ enter directly into the equations. This is because of the differing roles of electrons and ions in this case. The electrons only provide conductivity and since this is anisotropic the parameter p as well as ω_p plays a part. The ions however provide the inertia of the medium so that any movement in ions appears as a $v\times H$ force rather than conductivity due to ions.

This is to be compared (say) with the behaviour of a lightly ionized dense ternary gas. In such a gas the gyro and plasma frequencies of both electrons and ions influence the conductivity and therefore will appear in the dispersion relationship. The neutral particles will provide the inertia.

Acknowledgements—The author thanks Mr. T. Pearcey who performed the numerical calculations on C.S.I.R.A.C. in the Computations Laboratories at the University of Melbourne. The co-operation of Dr. F. Hirst and staff of the laboratories is gratefully acknowledged. Thanks for continued interest and criticism is extended to Dr. F. Jacka, Chief Physicist of this Division.

REFERENCES

1. K. D. Cole, *Planet. Space Sci.* **1**, 319 (1959).
2. J. H. Piddington, *Mon. Not. R. Astr. Soc.* **114**, 638 (1954).
3. S. K. Mitra, *The Upper Atmosphere*. The Asiatic Society, Calcutta (1952).
4. E. Astrom, *Ark. Fys.* **2**, 443 (1950).
5. J. H. Piddington, *Mon. Not. R. Astr. Soc.* **115**, 671 (1955).
6. K. D. Cole, *Aust. J. Phys.* **3**, 484 (1960).
7. K. D. Cole, *Nature* **186**, 874 (1960).
8. H. Alfven, *Cosmical Electrodynamics*. Clarendon Press, Oxford (1950).

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 299-306. Printed in Great Britain.

UPPER ATMOSPHERIC TEMPERATURES FROM DOPPLER LINE WIDTHS—I. SOME PRELIMINARY MEASUREMENTS ON 0I 5577 Å IN AURORA

J. A. NILSON* and G. G. SHEPHERD
Physics Department, University of Saskatchewan, Saskatoon, Canada

(*Received* 2 *February* 1961)

Abstract—A brief description is given of a photo-electric Fabry–Perot interferometer constructed for auroral and airglow studies. The instrument has plates 10 cm in diameter and employs refractive index scanning. An account is given of a study of the etalon characteristics and the resulting effect on the instrumental profile and the measured line widths. Preliminary measurements on the 0I 5577 Å auroral line are presented. These are distributed over a temperature range of 220°K to 700°K, with about half of the values lying between 400°K and 500°K.

1. INTRODUCTION

Ever since Babcock's[1] original interferometric observation of the 0I 5577 Å line in the night airglow it has been realized that upper atmospheric temperatures could be deduced from measured line widths. However, the resolution of Babcock's Fabry–Perot inferometer was just sufficient to place an upper limit on the width of the line. Subsquent observations on the airglow[2] and the aurora[3] consisted primarily of wavelength measurements and consequently were made at too low an order to provide data on widths. It is only recently, by using high orders and dielectric reflective coatings on the plates, that it has been possible to obtain line widths from the photographed fringes.[4–6] The scanning photoelectric Fabry–Perot interferometer has also been recently developed,[7, 8] and some measurements on the airglow and the aurora have already been made[9] with such an instrument.

The measurement of the line widths of forbidden transitions is particularly desirable. Because of the long lifetime of the excited state there is little doubt that the radiating atoms are in thermal equilibrium with the atmosphere and that the Doppler temperatures are in fact kinetic temperatures. The requirement that the natural line width be negligible is also well satisfied for a forbidden transition, and in addition, there is no hyperfine structure for O^{16} because of zero nuclear spin.

Most of the previous measurements have been on the nightglow, with the auroral observations being rather limited. Because of this it was decided to construct a Fabry–Perot interferometer for a continued observing program at Saskatoon, where aurora frequently occurs. A low-order Fabry–Perot spectrometer with a resolution of about 2·5 Å had already been used here for sodium twilight observations.[10] It should be of value to compare Doppler temperatures with other simultaneously measured spectroscopic temperatures, such as rotational temperatures from N_2^+ bands.

Although the theoretical basis for the derivation of kinetic temperature from the line profile is simple the same cannot be said for the experimental part of the problem. As the resolution of the apparatus is increased its light gathering power is decreased. Hence, it is not possible to have an instrumental profile width that is negligible in comparison with the 0I line width. In practice, the instrumental width is usually roughly equal to, or at best somewhat smaller than, the actual line width. It is obvious that rather careful procedures in measurement

* Present address: Department of Electrical Engineering, Imperial College, London.

and analysis must be followed to obtain reliable temperatures.

The relative advantages of the photoelectric and photographic instruments may be briefly considered. The advantages of the photoelectric detector are its great sensitivity and its linearity. It is true that a more complicated apparatus must be used since scanning is required, but it has been demonstrated[7, 8] that this is quite feasible. Scanning has the additional advantage that a spectrum is obtained directly, whereas with the photographic method considerable analysis of the photographed fringes is required. The instrument to be described here has obtained spectra photoelectrically in a few seconds even for faint aurora. Photographic instruments require exposure times measured in hours, which makes it impossible to measure the temperature for a single auroral form. For the nightglow radiations the source brightness is near the threshold of the photoelectric instrument and so the photographic method is most advantageously used here. However, by resorting to photon counting,[9] perhaps in conjunction with a memory unit,[11] the photoelectric instrument can be used here also with its accompanying advantages.

2. INSTRUMENTAL

The instrument is in general similar to that described by Karandikar[12, 13] and by Armstrong[14]. Only the significant points of difference will be described here.

Since refractive index scanning is used, the etalon is contained in a pressure cell. This is shown in Fig. 1. The Fabry–Perot plates (F) are separated by a 1 cm spacer (G) made from three lengths of invar tubing soldered in a brass ring, and are contained in a steel holder (E). Because the optical axis is vertical very little force is required on the adjusting screws to keep the etalon stable. The pressure cell (A) is of cast aluminum with 1 in. walls. The etalon holder touches the cell at only three points and is thus well isolated from temperature variations and mechanical flexure of the pressure cell. The cell is closed above and below by 12·5 cm diameter plate-glass windows. Shafts with "O" ring seals (not shown in the Fig.) in cover plate (D) allow the Fabry–Perot plates to be adjusted without opening the cell. The outer wall of the cell is wound with both heating wire and temperature sensing resistance wire for the controller.

The method of pressure scanning is that of supersonic flow through a capillary.[15] Freon-12 vapour in equilibrium with the liquid at room temperature is connected to one side of the capillary and the other side supplies the chamber with a nearly constant mass flow. The time variation of pressure has a linearity of 1 per cent or better over a range of 0 to 58 cm Hg, causing a scan over twenty-four orders. For auroral observations the scanning is only over a few orders so that the linearity is excellent. An automatic cycling arrangement is employed, using magnetic valves and pressure limit switches.

The general schematic arrangement of the entire system is shown in Fig. 2. All the mechanical components are mounted on a base plate which is supported in a Dexion framework by vibration mounts. The objective lens is an f/10 telescope objective of approximately 1 m focal length, and the circular aperture at its focus has a diameter of about 3 mm.

Below the aperture are an interference filter (not shown), a field lens and an E.M.I. 9502 photomultiplier cooled by a commercial refrigeration unit. The interference filter has a half-width of about 20Å with a transmission of 70 per cent, and is useful in reducing the white light background arising from scattered city lights and moonlight. The preamplifier is of the type described by Hunten[16] and the output is recorded on a Varian G–10 chart recorder.

A plane mirror system is placed above an opening in the roof immediately over the instrument. Since a small field of view is desired no lenses are placed above the etalon. The field of view is thus determined by the aperture and is about 0·2°.

3. THE ETALON CHARACTERISTICS

Chabbal[17] has shown that $Y(\sigma)$, the output signal from a Fabry–Perot spectrometer, can be considered as the convolution of three other functions. (Y is called the recorded function and σ denotes wavenumber.) These are the source function B, the etalon function E, and the aperture function F. In addition, E may be considered as the convolution of an Airy function A (for a perfect etalon) and a function D representing the deviation of the etalon from perfection.

The contours of the plates were observed as described by Chabbal[18], by viewing them in collimated light at low order. With the etalon in the pressure cell it is possible to change the optical path length by a known amount. This allows one to draw successive lines of constant plate separation on the etalon. Nine-layer zinc sulphide and cryolite coatings were used here with a plate separation of 0·03 cm. With the etalon in good adjustment the contours were observed to be roughly concentric circles, having at the centre of the plates a height of 0·035 λ with respect to the edges. Superimposed on this at the centre of the plates was a small hump 2·5 cm in diameter with a height of 0·015 λ. This means an extreme difference of 0·05 λ or 0·1 order. It was not ascertained to what extent the contours belonged to the dielectric layers rather than to the quartz.

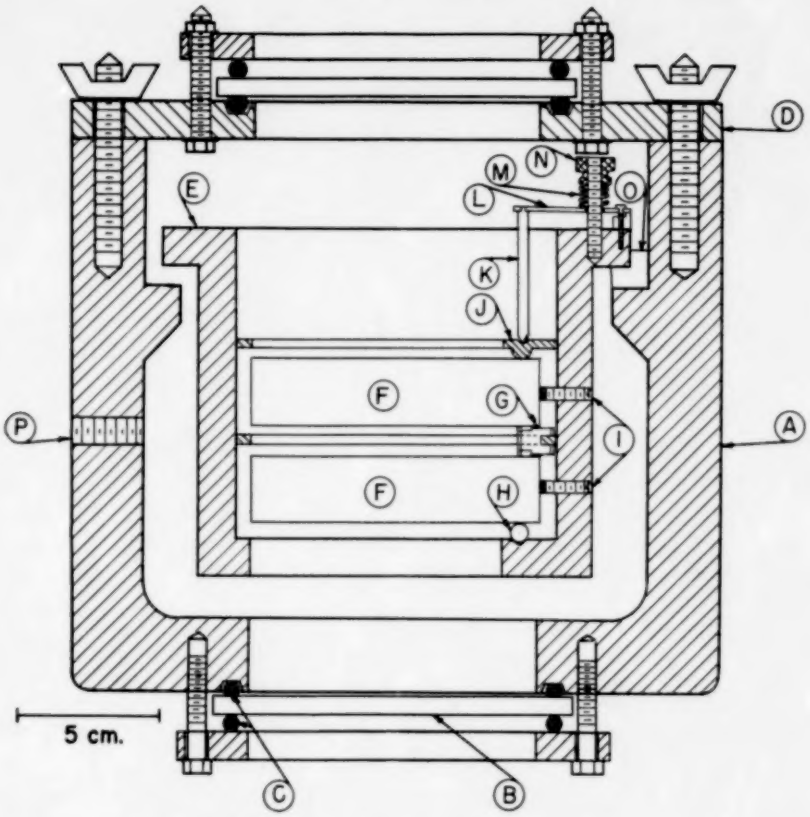


Fig. 1. Cross-sectional view of pressure cell and etalon. A, pressure cell; B, windows; C, "O" rings; D, cover plate; E, etalon holder; F, Fabry–Perot plates, G, invar spacer; H, steel ball; I, positioning screws; J, force ring; K, force rod; L, leaf spring, M, coil spring; N, adjustment nut; O, polystyrene support; P, pressure inlet.

When a spectral line is scanned at low order with a small aperture, the widths of B and F are negligible and Y becomes identical with E. An example of such a recorded function for 9-layer coatings is shown as the solid curve (a) in Fig. 3. The half-width is 0·030 order, corresponding to a finesse (ratio of free spectral range to function half-width) of 33. The width at the base is about 0·1 order, in agreement with the contours. The shape is roughly Gaussian, but is rather asymmetrical, with a lobe on one wing. The dotted curve (b) in Fig. 3 shows E obtained with the central 2·5 cm portion of the etalon masked, and is much more symmetrical. Evidently this portion of the plates can be identified with the lobe on the etalon function and the hump seen in the contours.

In addition to the large-scale deviations from flatness there are micro-deviations left after polishing. The effects of these were observed by masking all but a 1 cm circle of a flat portion of the etalon and tracing out $Y=E$ as before. The result is shown as the broken curve (c) in Fig. 3. The associated finesse is $N_E=51$ compared to $N_E=33$ for the whole plates. Considering E as the convolution of A and D, taking $N_A=105$ for nine-layer dielectric coatings[19] and assuming D to be Gaussian one obtains $N_D=41$ and 85 respectively for the whole etalon and for the micro-deviations.

To obtain the maximum luminosity at a given resolution the half-widths of A and D should be roughly equal. Therefore, 7-layer coatings were used for the auroral observations. The etalon function under these conditions is shown as the solid curve (a) in Fig. 4. E was also obtained for a small flat portion of the plates as before, and was almost identical with

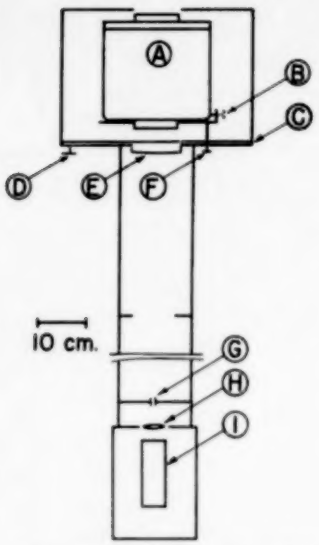


Fig. 2. General schematic side view of the Fabry–Perot spectrometer. A, pressure cell; B, positioning screws (3); C, baseplate; D, vibration mounts (4); E, objective lens; F, tilting screws (3); G, aperture; H, field lens; I, photomultiplier.

that for the whole etalon. This indicates that A and D have comparable widths and that masking is unnecessary for 7 layers. The half-width is 0·048 orders, giving $N_F=21$. A large increase in the product of luminosity and resolution was observed with the 7 layers.

The importance of good adjustment is also shown in Fig. 4. The dotted curve (b) was obtained with one

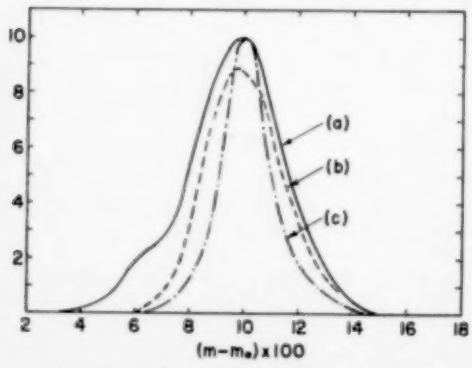


Fig. 3. Traces of the etalon function obtained at low order using 9-layer coatings. Solid curve (a), etalon unmasked; dotted curve (b), central 2·5 cm masked; broken curve (c), normalized function representing micro-deviations.

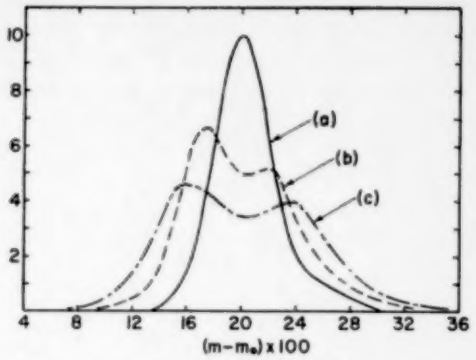


Fig. 4. Traces of the etalon function obtained at low order using 7-layer coatings. Solid curve (a), etalon in adjustment; dotted curve (b), one spacer compressed 0·042 orders; broken curve (c), one spacer compressed 0·091 orders.

spacer compressed by 0·042 orders, and the broken curve (c) by compressing the same spacer 0·091 orders. It is seen that the etalon function can become asymmetrical and even develop a central minimum.

The choice of etalon spacing requires a compromise between high resolution and a large free spectral range between orders. A good criterion for resolution is the ratio b/y, where b is the half-width of B (for the OI auroral line) and y is the half-width of Y. It was considered that $b/y=0·8$ would be desirable and $b/y=0·9$ would be excellent. Calculations indicated that with 7-layer coatings the value 0·8 could be achieved with a 1 cm spacing, and 0·9 with a 2 cm spacing. The value 0·8 could also be obtained with 5-layer coatings and a 2 cm spacing. It was decided to use a 1 cm spacer with 7-layer coatings to give just adequate resolution with the largest possible free spectral range. This completely eliminates overlapping orders and allows the base line to be established in the presence of dark current and white light.

4. ANALYSIS OF RECORDED PROFILES

There are several methods which may be used to obtain the actual profile (and thus the auroral temperature) from the recorded one. The first method is to assume that the etalon function obtained at low order can be used as the etalon function at high order when the observations are made. This function is then convolved with F, the rectangular function for the circular aperture, and finally with B, the Doppler profile for OI corresponding to some temperature T^1, giving Y, the recorded function. T^1 is adjusted until the calculated Y matches that produced by the instrument for the aurora. T^1 is then equal

to T for the aurora. This iterative calculation must be done numerically, since E is generally not an analytical function. In a method developed by Krebs and Sauer[20], and used, for example, by Armstrong[9], it is assumed that E is an Airy function and the convolution of this with a Doppler profile is expressed analytically as a series. This simplifies the calculation considerably. In actual practice the etalon function may not be constant, even for a carefully constructed instrument. Thus e, the half-width of E, must be found for the time of observation by recording a profile for a line of known half-width. This clearly doubles the amount of calculation.

A more elegant method, requiring no assumptions about the shape of the etalon function, does exist. If the instrumental profile is represented by some function W the recorded function will be the convolution of W with B. Then the Fourier transform of Y is equal to the product of the Fourier transforms of W and B. W is obtained by dividing the Fourier transform of Y by that of B for a line of known half-width, followed by an inverse transformation. This is done numerically with an electronic digital computer. Once W is known, B for 0I 5577 Å can be obtained by a similar calculation.

Because large numbers of observations are anticipated, a third method, requiring only a measurement of the half-width, was considered. Graphs given by Chabbal [7, 17] allow one to find the half-width of a function, which when convolved with a function of known half-width give a third function whose half-width is known. These graphs include cases where each of the convolved functions is either Airy, Gaussian or rectangular. In the present problem, two convolutions must be taken in order, written $Y=B\bigstar E\bigstar F$. This can be done with these graphs only if the product of the first convolution is also one of the three functions. In Section 3 it was shown that E is very close to being Gaussian for this etalon, for both 7-layer and 9-layer coatings. Since B is Gaussian, $B\bigstar E$ is also Gaussian. (For this convolution the graphs need not even be used, since $b^2+e^2=g^2$, where b, e and g are the half-widths of B, E and $G=B\bigstar E$, respectively.) Thus this method appeared very promising for this instrument. The convolution of F and G can then be carried out graphically.

The line-width reference used during the present observations was the *Kr* 5570 Å line, which is conveniently close to 5577 Å. The Doppler half-width for *Kr* for this lamp was unknown. But applying the e values of Section 3 to the narrowest *Kr* profiles obtained at high order (using the method outlined above) gave a *Kr* half-width of 35 mk. (1 millikaiser = 1/1000 wavenumber). During auroral observations e values were usually larger than this. They were determined for a particular set of auroral observations from *Kr* profiles recorded at the same time.

In order to confirm the reliability of the method proposed above (which will be referred to as the half-widths method) one set of profiles was calculated using numerical Fourier transforms. This was done using an LGP–30 Royal McBee digital computer. The results are shown in Fig. 5. Curve (c) is the observed *Kr* profile and (d) is its assumed profile, Gaussian with a half-width of 35 mk. The instrumental profile W, curve (e), was obtained by dividing the Fourier transform of (c) by that of (d) and taking the inverse transform. Dividing the Fourier transform of the observed 0I profile (a) by that of W (e), and taking the inverse transform, the derived 0I profile shown as curve (b) was

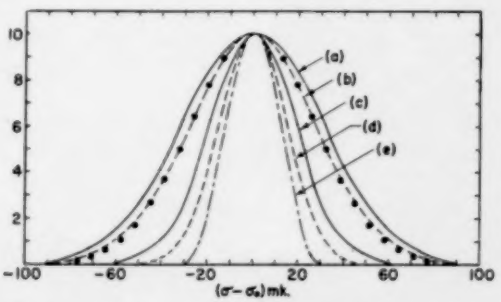


Fig. 5. Results of numerical Fourier transform calculations. (a), observed 0I 5577 Å profile; (b), 0I 5577 Å profile derived from the calculations; (c) observed *Kr* profile; (d) assumed *Kr* profile of half-width 35 mk; (e) instrumental profile derived from (c) and (d) and used to obtain (b) and (a). All profiles were assumed symmetrical for the calculations. The solid circles are for a Gaussian having the same half-width as (b).

obtained. The difference between (a) and (b) is surprisingly small, considering that W has a half-width more than a third that of the observed profile. The derived profile has a half-width of 64·5 mk compared with 66·4 mk using the half-widths method, which is considered satisfactory agreement. A further check on the method was obtained by considering W the convolution of F and E. Using $f=20$ mk a value of e of 24·0 mk was obtained, compared with 24·9 mk given directly by the half-widths method. The dots shown in Fig. 5 are for a Gaussian of 64·5 mk half-width, and show that the derived 0I line is very closely Gaussian. This is a confirmation of the validity of the interpretation as well as of the experimental procedures.

5. RESULTS

Before the study of etalon characteristics described in Section 3 had been completed some auroral observations were made using 9-layer dielectric coatings. Since the etalon transmission would be very low if Na/Nd were large, and Na is approximately 100, this confirmed the high quality of the plates. These observations were made on 31 May, 3 June and 5 June 1960. Subsequently, with 7-layer coatings, aurora was observed on 12 July and 13 July 1960. The auroral profiles were carefully selected from those available. The only ones chosen were symmetrical and in a sequence of several having the same peak intensity indicating that the auroral intensity had not changed during the scans. Those in a sequence were averaged to give one width and one temperature.

The results are summarized in Table 1, where the number in a given temperature range is shown for various intervals. It is seen that most of the temperatures are between 300°K and 500°K. The lowest obtained was 220°K and the highest was 700°K. Examples of recordings obtained in 40 sec using 9-layer coatings are given in Fig. 6. The time required to record one fringe is thus 10 sec or less. The dotted trace was obtained from an arc to the south, and gave a temperature of 405°K. The solid trace was made on a mixture of rays and bands to the east, and gave a temperature of 700°K. The broken trace is for the krypton lamp. All three traces were obtained within a few minutes of each other. Even without normalizing them it is obvious that the half-widths are different, confirming that the actual line width makes a substantial contribution to the recorded line width. The noise level for the runs shown is somewhat higher than usual. It is seen that a well-defined base-line is established between orders.

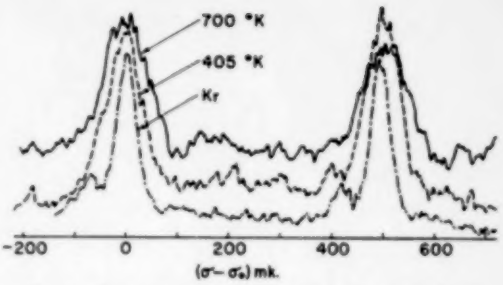


Fig. 6. Superimposed examples of recordings obtained for 0I 5577 Å in aurora, using 9-layer coatings. The recording time for one record is 40 sec.

The probable errors of measurement vary greatly depending on the auroral intensity and the resulting signal to noise ratio. They will not be assessed here. The systematic errors are more important and an attempt was made to place extreme limits on them in the following way. It is impossible for the etalon finesse to exceed

Table 1. Summary of Doppler temperatures from 0I 5577 Å in aurora

Temperature range (°K)	Number of observations	Auroral type
201–300	3	arcs
301–400	8	arcs, glow and diffuse
401–500	12	arcs, band, ray, glow, diffuse
501–600	2	glows
601–700	1	mixture rays and band

Table 2. Extreme limits for some Doppler temperatures

Number of layers	e (mk)	y (mk)	b/y	Lower limit (°K)	Assigned temp. (°K)	Upper limit (°K)
9	60	78	0·61	154	222	558
9	38	68	0·80	215	283	409
7	25	66	0·89	272	344	360
7	25	72	0·91	354	422	444
7	32	84	0·90	503	570	644
9	47	98	0·86	630	700	897

25 for 7-layer coatings or 40 for 9-layer coatings. Using the appropriate one of these values gives the maximum temperature that can be derived from a given profile. Also, the actual *Kr* half-width can not be less than its Doppler value at room temperature of 25 mk. Using this in place of the 35 mk adopted value gives the lowest temperature from a given auroral profile. In Table 2 six of the temperatures obtained are listed, along with the upper and lower limits of temperature so obtained. Along with these are given the number of dielectric layers, the half-width e of the etalon function, the half-width y of the recorded auroral profile and the ratio b/y where b is the derived auroral half-width. Even if the different temperatures were in error by the extreme limits, which is very unlikely, the full range of measured temperatures cannot be accounted for as arising from systematic errors. The mean difference between assigned temperature and lower limit is 70°K, and between assigned temperature and upper limit, 135°K. Large differences for the upper limits naturally correspond to the cases of poor adjustment (large e). In particular, the first case combines low temperature with very poor adjustment and gives a difference of 336°K. The third case has a difference of only 16°K. It was at first surprising that smaller values of e are obtained with 7-layers than with 9. This is attributed partly to inferior adjustment during the 9-layer observations, since the etalon contours were not yet known. In addition, the external adjustment mechanisms had not yet been installed, and in some cases observations were made rather soon after assembling the etalon. It illustrates very well that extreme care must be taken if one is to benefit from using 9-layer coatings. The values of b/y vary considerably depending on temperature and etalon adjustment, but are as great as 0·9, so that the recorded profile can closely approach the actual one.

6. DISCUSSION

Although only 26 temperatures have been obtained, they cover a wide range, from 220°K to 700°K. Upper and lower limits have been placed on six of these values, and definitely show that the temperature variation is real. These results are in general agreement with those of Armstrong[9], who obtained one temperature in the 201–300°K range, four in the range 301–400 °K, one in the range 401–500°K and two in the range 501–600°K. The distribution is rather different from that obtained for rotational temperatures from N_2^+ bands by Shepherd and Hunten[21]. Their temperatures were divided into three groups; low (150–220°K), average (221–274°K), and high (275–465°K). The numbers of measurements for each group were 59, 64 and 46 respectively. By comparison, the corresponding numbers of Doppler temperatures are 0, 1 and 20, with five above 465°K. However, rotational temperatures from N_2 bands obtained by Montalbetti[22] give the numbers 0, 4 and 11, which more closely resemble those for the Doppler temperatures. It is possible that the differences between the N_2^+ rotational and Doppler results for Saskatoon are a result of observational selection. The Doppler temperatures were obtained during the summer on rather faint, quiet forms. The rotational temperatures were mostly obtained in the spring from more intense, active forms. This explanation is supported by Montalbetti's[22] results, which were obtained mostly from aurora of

intensity 1 and 2. It follows that the faint, diffuse forms have greater heights than the intense ones, as is generally believed.

The above results should be regarded as preliminary, since they are limited in number and the method of analysis has not yet been fully explored. Nevertheless they demonstrate that reliable Doppler temperatures can be obtained from aurora with scanning times of a few seconds. It is intended that a more definite value be obtained for the width of the *Kr* line used, by comparing it with a Hg^{198} lamp. The correct value is probably somewhere between 25 mk and 35 mk, so that the correct temperatures lie somewhere between the lower limits and the assigned temperatures. Much more work on the method of analysis is possible and this will be continued. It is intended that many more measurements be made on 0I 5577 Å, and that they be extended to 0I 6300 Å. With the narrow field of view it should be possible to measure temperature gradients in aurora with vertical extent. Simultaneous Doppler and N_2^+ rotational temperature measurements will also be considered.

Acknowledgements—This research was supported by a grant from the National Research Council of Canada. Some additional financial assistance from the Institute of Upper Atmospheric Physics, University of Saskatchewan, is gratefully acknowledged.

The authors wish to thank Dr. D. M. Hunten for supplying the photomultiplier detector and interference filter, and for making available the equipment for dielectric coating. Dr. H. P. Gush, University of Toronto, gave useful advice concerning the Fourier transform calculations. Mr. E. C. Turgeon assisted with the low-order measurements.

REFERENCES

1. H. D. Babcock, *Astrophys. J.* **57,** 209 (1923).
2. J. Dufay, J. Cabannes and J. Gauzit, *Astronomie* **56,** 149 (1942).
3. L. Vegard, *Phil. Mag.* **24,** 588 (1937).
4. D. Q. Wark and J. M. Stone, *Nature Lond.* **175,** 254 (1955).
5. J. Cabannes and J. Dufay, *The Airglow and the Aurorae.* (Edited by E. B. Armstrong and A. Dalgarno) Pergamon Press, London and New York (1956).
6. D. Q. Wark, *Astrophys. J.* **131,** 491 (1960).
7. R. Chabbal, *Rev. Opt. (théor. instrum.)* **37,** 49 (1958).
8. P. Jacquinot, *Rep. Progr. Phys.* **23,** 267 (1960).
9. E. B. Armstrong, *J. Atmos. Terr. Phys.* **13,** 205 (1959).
10. G. G. Shepherd, *Canad. J. Phys.* **38,** 1560 (1960).
11. D. M. Hunten, *Canad. J. Phys.* **38,** 346 (1960).
12. R. V. Karandikar, *Rev. Opt. (théor. instrum.)* **35,** 414 (1956).
13. R. V. Karandikar, Final Report, Air Force Cambridge Research Center, Bedford, Mass. (1959).
14. E. B. Armstrong, *J. Phys. Radium* **19,** 358 (1958).
15. H. D. Rank, *J. Opt. Soc. Amer.* **44,** 5 (1954).
16. D. M. Hunten, *Canad. J. Phys.* **31,** 681 (1953).
17. R. Chabbal, *J. Rech.* **24,** 138 (1953).
18. R. Chabbal, *J. Phys. Radium* **19,** 295 (1958).
19. P. Giacomo, *C.R. Acad. Sci., Paris* **235,** 1627 (1952).
20. K. Krebs and A. Sauer, *Ann. Phys., Lpz.* **13,** 359 (1953).
21. G. G. Shepherd and D. M. Hunten, *J. Atmos. Terr. Phys.* **6,** 328 (1955).
22. R. Montalbetti, *Canad. J. Phys.* **35,** 831 (1957).

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 307-313. Printed in Great Britain.

LUNAR TEMPERATURE MEASUREMENTS AT 3200 MC/S

W. J. MEDD and N. W. BROTEN
Radio and Electrical Engineering Division, National Research Council, Ottawa, Canada

(*Received* 6 *February* 1961)

Abstract—The measured effective black-body temperature at the lunar equator is 220°K, with an estimated probable error of 5 per cent. No variation with lunar phase was detected; if present, and obscured by the random scatter of the observations, it is less than 5 per cent at this frequency.

1. INTRODUCTION

Measurements of the temperature of the moon at infrared wavelengths have been available for many years; we may refer particularly to the classic investigation of Pettit and Nicholson[1], Pettit[2] and the theoretical studies by Wesselink[3]. In the microwave region ($\lambda = 1{\cdot}25$ cm) the first measurement was made by Dicke and Beringer[4] in 1946, but the work of Piddington and Minnett[5] in 1949 was of special interest, inasmuch as a series of measurements of lunar temperature was taken as a function of lunar phase, and the results were in marked contrast with those in the infrared. These measurements were subjected to a careful analysis by the authors, Pawsey and Bracewell[6], Jaeger[7] and Troitskii[8, 9], primarily with a view to deriving additional information on the material comprising the lunar surface.

Since 1949, about a dozen separate investigations in the microwave region have been reported. These are summarized in Table 1. The equation

$$T_m = T'_m - T' \cos(\omega t - \psi) \qquad (1)$$

may be used to specify (to a first approximation) the temperature of a point on the lunar surface, restricted in the present instance to the lunar equator. T'_m is the average temperature of this point over the lunar day of period $P = \omega/2\pi$, and ψ is the phase displacement in degrees of the minimum value of T_m from the time ($t = 0$) of new moon. The average brightness temperature is given by $T'_m(1-R)$, where R is the power reflectivity of the lunar surface. Values of $T'_m(1-R)$, T'/T'_m in percentage and ψ are listed. It will be noted that there is rather general agreement on the value of ψ, and that further information has been obtained on the modulation term (T'/T'_m) as a function of frequency. In this latter respect, the value given by Akabane[10] is anomalous, and a further measurement in the neighbourhood of 3 kMc/s is therefore desirable.

The absolute value of the lunar temperature is of theoretical importance. There is no reason to expect a variation of T'_m with frequency, and the individual measurements as given in Table 1 should be directly comparable. There are, however, certain difficulties which must be pointed out. An attempt has been made to reduce the results to a common denominator, equatorial brightness temperature, but in some cases it is not known whether the value reported is intended to specify the equatorial temperature, or an average over the entire lunar surface; the difference may be some 5 per cent or less. Piddington and Minnett[5] used a correction of 10 per cent in consideration of the power reflectivity of the lunar surface, but for several of the values in Table 1 it is not known whether such a correction has, or has not, been applied. Finally, it is very difficult to evaluate the accuracy of the individual data. In the original reports the descriptions of experimental technique are usually brief, if available at all. The statements of accuracy imply that it is unlikely (if not highly unlikely) that the over-all error exceeds some 5 to 20 per cent in different cases, but it seems that at least some of these estimates are too optimistic. Considering the over-all picture, and omitting a few extremes, it will be

noted that the values range from about 150°K to 270°K. This is more typical of what would be expected if the individual experiments were performed to an accuracy of about ±10 per cent probable error. An examination of the theoretical models of Jaeger[7] indicates that a determination of T_m to an accuracy of ±2 per cent probable error would be useful; at the present state of the art, this is perhaps possible, but extremely difficult.

2. OBSERVATIONS AND DETERMINATIONS OF VARIOUS PARAMETERS

(i) *General*

A description of the equipment and techniques of calibration has been given elsewhere[11], but will be reviewed briefly at this point.

The observations were taken during 1959 at Ottawa, Canada, with a 10-foot diameter parabolic reflector and a radiometer employing travelling-wave-tube amplifiers. A bandwidth of 450 Mc/s was used, centered at 3200 Mc/s.

(ii) *Required parameters*

T_m is related to the effective antenna temperature, T_A, by the following equation:

$$T_A = \frac{G_0 \gamma}{4\pi} [T_m \int (1-R)\, t(\alpha, \beta)\, g(\alpha, \beta)\, d\alpha d\beta] + C \qquad (2)$$

where G_0 is the maximum power gain of the antenna,
γ the transmission coefficient for the terrestrial atmosphere,
t the normalized spatial temperature distribution,
g the normalized power gain function of the antenna,
and C that portion of the antenna temperature due to contributions from the "background", specifically the remaining sky, the earth, and the terrestrial atmosphere.

The origin of the co-ordinate system is taken as the centre of the projected lunar disk.

(iii) *Antenna temperature*

The measurement of T_A is indicated in Fig. 1, which shows a drift-curve or meridian transit

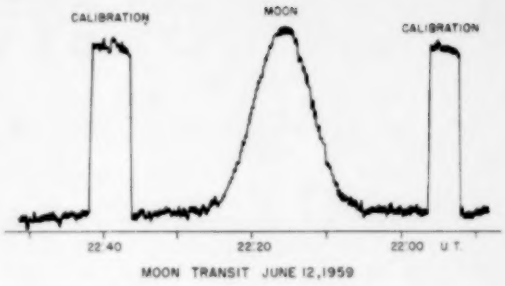


Fig. 1.

of the moon. The calibration marks were obtained from an argon discharge tube, which had previously been calibrated against a thermal load or heated resistive element. The probable error of the magnitude of the calibration signals is about ±3 per cent. The base-line was established prior to transit by adding to the antenna feed-line, through a directional coupler, sufficient power from an argon discharge tube to obtain a null reading on the recorder. (The receiver is switched at 30 c/s between the antenna and a resistive termination at ambient temperature.) Throughout the observing period the background radiation, C, remains present and essentially constant, the calibration signals being injected into the antenna feed lines. An explicit determination of C is therefore unnecessary and we obtain at the maximum of the drift curve an effective temperature,

$$T_{AM} = \frac{1}{4\pi} G_0 \gamma T_m \int (1-R)\; t\, g\, d\alpha\, d\beta \qquad (3)$$

Aside from records spoiled by interference, bad weather, etc., seventy-six observations similar to that shown in Fig. 1 were obtained over the course of a year.

(iv) *Antenna gain*

G_0 was found to be 3805 ±2 per cent (estimated probable error). This value was obtained by an extended series of measurements using the sun as a signal source and a 4 ft by 3 ft aperture horn as a comparison antenna. The calculated gain of the horn and the measurements of solar flux with the two antennas then suffice to establish G_0, the gain of the parabolic reflector.

(v) *Antenna pattern or gain function*

The normalized power gain function of the antenna was found from observations of the sun and point sources. Inasmuch as the antenna pattern is about 140 min of arc to the half-power points, compared with some 32 min for the diameter of the moon, it is found that $g(\alpha, \beta)$ is almost constant over the lunar disk. For convenience in the calculations we put $\alpha = rx$ and $\beta = ry$, where r is the semi-diameter of the moon, and the gain function is then approximated very closely by the expression $1 - A(x^2 + y^2)$. A is determined from the antenna pattern (using an average value of r) as 0·032. The error involved is negligible.

(vi) *Atmospheric attenuation*

At 3200 Mc/s γ is close to unity and introduces only a small correction. As a function of zenith angle z, we have $\gamma \doteqdot (-0{\cdot}037 \sec z)\, db$. The attenuation at the zenith of $0{\cdot}037\, db$ is derived from calculations, but is consistent with experimental evidence. The error involved is negligible.

(vii) *Reflectivity*

Piddington and Minnett[5] originally used a value for $(1 - R)$ of 0·9, while Gibson[12] used 0·85. These are reasonable figures based on a dielectric constant (ε) of about 5, typical of terrestrial rock. However, there is recent evidence from radar studies of the moon, as interpreted by Senior and Siegel[13] that the value of $(1 - R)$ should be close to one. Hughes[14] places the value somewhere between 0·9 and 0·99, whereas Fricker *et al.*[15] have found 0·98 from radar studies at 400 Mc/s. It would appear that the matter can be suitably deferred by calculating a brightness temperature, $T_m(1 - R)$. This will be done, but there is a further complication in that $(1 - R)$ is a function of lunar co-ordinates. We may write $(1 - R) = (1 - R_\perp)\, \Gamma(\varepsilon, \eta)$, where $\eta^2 = \alpha^2 + \beta^2$, and $R_\perp$ is the power reflectivity at normal incidence. The effect of the function $\Gamma(\varepsilon, \eta)$ has been discussed in Ref. (8). The value is quite constant and close to one over the central regions, but may produce an appreciable limb-darkening unless R is extremely low (ε close to 1), or unless the function is modified appreciably by the roughness of the lunar surface.

(viii) *Temperature distribution*

Measurements of the lunar temperature as a function of longitude have been made in the infrared[1]. Assuming circular symmetry and ignoring local variations due to surface roughness it was found that $T \doteqdot T_{ss} \cos^{1/6} \rho$, where T_{ss} is the temperature of the subsolar point. At microwave frequencies, where phase lags occur over the lunar day, the temperature distribution as viewed at any particular time may be a rather unwieldy function of the lunar co-ordinates. However, it is possible to make certain approximations. First, the geometry is simplified by assuming that the subsolar point follows the lunar equator, and that the origin of our co-ordinate system (α, β) also falls on the lunar equator. It is expected that the error so introduced is negligible, which is verified after the calculations are complete. The distribution of temperature with lunar longitude, θ, is related to the variation of temperature with lunar phase. The mathematical exposition is greatly simplified by tentatively assuming that $T(\theta, \phi)$ is constant as a function of θ for any particular observation, and thus calculating for observations at different lunar phases a running average with respect to $T(\theta)$. If $T(\theta)$ follows a sinusoidal distribution we then find by this method a modulation term which is about 0·6 times the true phase-temperature variation.

It remains then to specify the lunar temperature as a function of the latitude ϕ. A direct measurement of this function, as well as $T(\theta)$ and ψ, is of course highly desirable, but at 3000 Mc/s an antenna of some 600 ft diameter would be required in order to obtain the necessary resolution. Indirect evidence is available from measurements in the infrared[1, 16], and from recent surveys at 4·3 mm by Coates[17] and at 8 mm by Amenitskii *et al.*[18]. We shall take $f(\phi) = \cos^{1/4} \phi$, which appears to be a reasonable estimate. This has been widely used in the literature, although

$$f(\phi) = 0{\cdot}32 + 0{\cdot}68 \cos^{1/2} \phi,$$

which is not appreciably different, has also

been used. (The infrared measurements give 120°K as the temperature of the dark side of the moon, and 374°K as the temperature of the subsolar point; the ratio is 0·32.) The error involved has been estimated by repeating the calculations with somewhat more uniform and more tapered distributions of temperature; considering $t(\alpha, \beta)$ as introducing a multiplicative factor in the final results, it has been assigned an estimated probable error of about 2 or 3 per cent.

3. CALCULATIONS

Assuming first that $\Gamma(\varepsilon, \eta) = 1$, and that $1 - R = 1 - R_\perp$, we have from equation (3) that

$$T_{AM} = \frac{1}{\pi} G_0 \gamma T_m (1 - R_\perp) \int_{\beta=0}^{r} \int_{\alpha=0}^{\sqrt{(r^2-\beta^2)}} t(\alpha, \beta)\, g(\alpha, \beta)\, d\alpha\, d\beta \qquad (4)$$

where r is the semi-diameter of the moon. A change to (x, y) co-ordinates, where $\alpha = rx$ and $\beta = ry$, is convenient. As previously discussed, we put

$$g(x, y) = 1 - A(x^2 + y^2)$$

and

$$t(\theta, \phi) = \cos^{1/4}\phi, \text{ or } t(x, y) = (1 - y^2)^{1/8}$$

whence

$$T_{AM} = \frac{1}{\pi} G_0 \gamma T_m (1 - R_\perp) \int_{y=0}^{1} \int_{x=0}^{\sqrt{(1-y^2)}} (1 - y^2)^{1/8} (1 - Ax^2 - Ay^2)\, dx dy.$$

Substituting $G_0 = 3805$, $A = 0{\cdot}032$ and performing the integration it is found that

$$T_m (1 - R_\perp) = \frac{13230}{\gamma r^2} T_{AM} \qquad (5)$$

where r is now given in minutes of arc rather than radians, and $\gamma = (-0{\cdot}037 \sec z)\, db$.

Let us now consider the effect of including some calculated values for $\Gamma(\varepsilon, \eta)$, which has been assumed equal to one in equation (5). R, an average of the vertically and horizontally polarized components, is given by the expression

$$R = \frac{1}{2}\left\{\left(\frac{\varepsilon \cos p - q}{\varepsilon \cos p + q}\right)^2 + \left(\frac{\cos p - q}{\cos p + q}\right)^2\right\} \text{ where } q^2 = \varepsilon - \sin^2 p.$$

This function is not readily handled analytically in equation (4); however, numerical integrations have been performed with the following results:

ε	$1 - R_\perp$	$T_m (1 - R_\perp)$
1·1	0·9994	$\frac{13230}{\gamma r^2} T_{AM} \times 1{\cdot}01$
1·5	0·99	$\frac{13230}{\gamma r^2} T_{AM} \times 1{\cdot}035$
3·0	0·93	$\frac{13230}{\gamma r^2} T_{AM} \times 1{\cdot}07$
5·0	0·85	$\frac{13230}{\gamma r^2} T_{AM} \times 1{\cdot}07$

It appears that a multiplicative factor between 1·01 and 1·07 should be applied in equation (5), the particular value being arbitrary inasmuch as the precise value of ε is not known. We have used 1·02, a comparatively low value being assumed, inasmuch as the available evidence suggests a low value of ε, and as the roughness on the lunar surface may tend to diminish this limb-darkening effect. We have finally then that

$$T_m (1 - R_\perp) = \frac{13500}{\gamma r^2} T_{AM}.$$

4. RESULTS

(i) *Variation with lunar phase*

The observations were taken over a period of ten lunations. No additional information was obtained by plotting the lunations separately, and Fig. 2 is a composite plot of equatorial lunar temperature as a function of lunar phase. No functional variation is evident. It is possible, of course, that a small effect is obscured by the scatter of the observations. From a statistical analysis of the data and considering also the smoothing effect

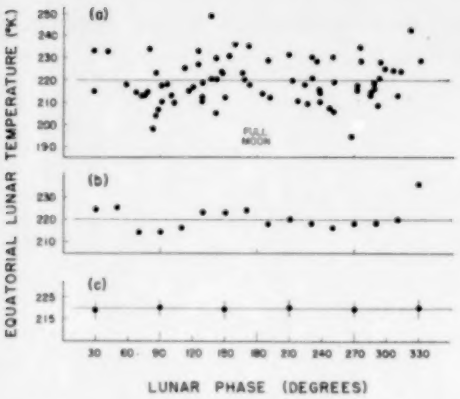


Fig. 2. Equatorial temperature as a function of lunar phase:
(a) Individual observations.
(b) Averages for 20-degree phase intervals.
(c) Running averages for 180-degree phase intervals. The vertical lines represent four times the standard deviation of the mean.

due to the width of the antenna beam it is found that any variation with lunar phase is likely (probability of 0·95) to be less in amplitude than 5 per cent.

(ii) *Average equatorial temperature*

The results of the seventy-six observations are shown plotted in histogram form in Fig. 3. The mean value of the equatorial lunar-brightness temperature is 220°K. The scatter of the observations is somewhat larger than expected; however, the probable error of the mean is less than 1 per cent and this is not significant when compared with other sources of error. The over-all probable error is about 5 per cent. This is, in the main, an estimated value and is not derived from statistical analysis. The principal sources of error are in G_0—2 per cent systematic; in $t(\alpha, \beta)$—3 per cent systematic; in $\Gamma(\varepsilon, \eta)$—less than 2 per cent systematic, and in the determination of T_{AM}—3 per cent experimental. Other parameters contribute individually less than 1 per cent.

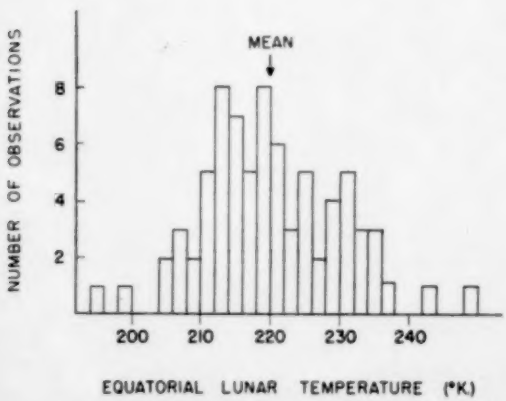


Fig. 3. Histogram of the temperature observations.

It is believed that this value is more reliable than most previous results. The present experiment suffers from the disadvantage that the calculations are comparatively critical to the uncertainties in $t(\alpha, \beta)$ and $\Gamma(\varepsilon, \eta)$. On the other hand, the antenna parameters have probably been more accurately specified, as the experimental method eliminates the necessity for measuring background temperature, and the gain function has been measured by reference to the calculated gain of a rectangular horn. With antennas of some 100 or 200 wavelengths diameter, as have been used in many of the previous experiments, the estimation of antenna parameters presents particular difficulties. For example, Gibson[12] has indicated that his value of $T'_m(1-R)$ (present nomenclature) of about 180°K may be raised by 10 per cent to allow for uncertainty in his antenna weighting factors, but Mezger and Strassl[19] have revised it by +43 per cent, and Giraud[20] by +25 per cent. It is a matter of speculation as to how many of the values listed in Table 1 are susceptible to similar differences of opinion.

With respect to the theoretical studies of Jaeger[7] a value of 220°K for the equatorial lunar temperature apparently leads to models of the lunar crust, which are in satisfactory agreement with the measurements in the infrared. It will be noted, however, that in consideration of the probable error of ±5 per cent the properties of the lunar surface are still not uniquely determined by the radio measurements alone.

Table 1. Values of average equatorial lunar brightness temperatures, $T'_m(1-R)$, percentage modulation, and phase displacement

Author	Reference	Frequency (kMc/s)	λ (cm)	Brightness temp. (°K)	T'/T'_m (%)	ψ (degrees)
(i) Mitchell *et al.*	21	40	0·75	150	17	
(ii) Salomonovich	22	37·5	0·8	197	16·3	40
(iii) Gibson	12	37·5	0·8	200	21·6	40
(iv) Hagen	23	35·5	0·85	150		
(v) Dicke and Beringer	4	24	1·25	270		
(vi) Piddington and Minnett	5	24	1·25	225	21	45
(vii) Zelinskaya *et al.*	24	18·4	1·63	224	16	40
(viii) Grebenkemper	25	13·6	2·2	220	5	
(ix) Troitskii and Zelinskaya	26	9·4	3·2	170	<7	
(x) Zelinskaya and Troitskii	27	9·4	3·2	183	<5	
(xi) Kaidanovskii *et al.*	28	9·4	3·2	120	<7·5	
(xii) Kaidanovskii *et al.*	28	3·0	10·0	117		
(xiii) Akabane	10	3·0	10·0	315	23	
(xiv) Mezger and Strassl	19	1·4	21·4	250	<2	
(xv) Westerhout	29	1·39	21·6	270		
(xvi) Denisse and Le Roux	30	0·9	33·4	208		
(xvii) Seeger *et al.*	30	0·4	75	195	<10	

Comments (in order of listing above). Comments by the respective authors are given in quotations.

(i) "This figure (150°K) should be considered tentative pending future investigations." (32 observations).

(ii) "The mean spread amounts to ±6°. The systematic errors . . . . do not exceed 10 per cent." (About 60 observations).

(iii) "The uncertainty of (the antenna-beam weighting factors) permits the measured curve . . . to be raised by some factor not greater than 10 per cent." The value listed (200°K) is accordingly about 10 per cent higher than that indicated in Gibson's paper. (About 30 observations).

(iv) "The average equivalent temperature . . . . measured on several days . . . . is 150°Kelvin ±40 per cent."

(v) The published value is 292° for the effective black-body temperature, averaged over the lunar disk. The following corrections have been applied:

1. To adjust for an incorrect value of the semi-diameter of the moon, minus 8 per cent (see Ref. 5).
2. To obtain an equatorial temperature, plus 8 per cent (an estimate only).
3. To obtain an equatorial temperature averaged over a lunar cycle, rather than the single day's observations (following Piddington and Minnett's curve of phase variation) minus 8 per cent.

"The probable error is believed to be about ±10 per cent."

(vi) "The absolute values are thought to be accurate within ±5 per cent." The published value of T_m is 249°, a 10 per cent correction for reflectivity having been applied. (31 observations).

(vii) "The r.m.s. error of the ordinate of the radio temperature curve amounts to about 5 per cent. However, the systematic error associated with inaccuracies in the determination of the antenna parameters and insufficient precision in calibration amounts to ±10 per cent to ±15 per cent." (Approximately 55 observations).

(viii) " . . . . observations were obtained for 5 days which happened to be spaced in the lunar cycle. The beam brightness temperature of the moon was found to be 200°K . . . . If (the spillover factor) is taken to be 0·1, the temperature values are increased 10 per cent; i.e. the mean brightness temperature is about 220°K at the centre of the lunar disk. . . . . the systematic error affecting the data as a whole probably is not greater than ±5 per cent."

(ix) " . . . . a chance error of the order of ±3 per cent in an individual determination. The systematic error due to inaccuracy of the calibration and determination of the parameters of the antenna is approximately 20 per cent, however." (70 days observations, 1952).

(x) Apparently the work of Troitskii and Zelinskaya was repeated in 1955 with greater accuracy. 23 observations were taken on 7 days and an accuracy of ±13° was placed on the value of 183°K.

(xi) Using $R=0·1$ the published value is (133 ±20) °K.

(xii) The published value is 130°K (±20 per cent). $R=0·1$ is probably used.

(xiii) "The errors . . . . are estimated to be about less than 50°K, which are introduced both from the estimation of the antenna gain and from the accidental errors of the experiments." (39 observations).

(xiv) "As opposed to the relative accuracy of ± 2 per cent, an estimation of the calibration errors for aerial temperature and aerial yields a value of ± 12 per cent for the absolute uncertainty of our results." (73 observations).

(xv) "The moon was observed on only one night . . . . near full moon. The brightness temperature of the disk is . . . . 270°K ± 60."

(xvi) Reported by Seeger *et al.* in Ref. 30. No details or statement of accuracy.

(xvii) The published value is 185°K, apparently an average over the lunar disk. The equatorial temperature would probably be about 6 per cent higher. No statement of accuracy is given, but the derivation is based on a previous study of Cassopeia A, where the error is estimated as 15 per cent.

REFERENCES

1. E. Pettit and S. B. Nicholson, *Astrophys. J.* **71,** 102 (1930).
2. E. Pettit, *Astrophys. J.* **81,** 17 (1935).
3. A. J. Wesselink, *Bull. Astro. Insts. Netherlds.* **10,** 351 (1948).
4. R. H. Dicke and R. Beringer, *Astrophys. J.* **103,** 375 (1946).
5. J. H. Piddington and H. C. Minnett, *Austral. J. Sci. Res.* Series A, **2,** 63 (1949).
6. J. L. Pawsey and R. N. Bracewell, *Radio Astronomy* p. 280. Clarendon Press, Oxford (1955).
7. J. C. Jaeger, *Austral. J. Phys.* **6,** 10 (1953).
8. V. S. Troitskii, *Astr. J. Moscow* **31,** 511 (1954).
9. V. S. Troitskii, *Trans. 5th Congr. on the Questions of Cosmogony, March,* 1955, p. 325. Akademia Nauk S.S.S.R. (1956).
10. K. Akabane, *Proc. Japan. Acad.* **31,** 161 (1955).
11. N. W. Broten and W. J. Medd, *Astrophys. J.* **132,** 279 (1960).
12. J. E. Gibson, *Proc. Inst. Radio Engrs. N.Y.* **46,** 280 (1958).
13. T. B. A. Senior and K. M. Siegel, *J. Res. Nat. Bur. Stand.* **64D,** 217 (1960).
14. V. A. Hughes, *Nature, Lond.* **186,** 873 (1960).
15. S. J. Fricker, R. P. Ingalls, W. C. Mason, M. L. Stone and D. W. Swift, *J. Res. Nat. Bur. Stands.* **64D,** 455 (1960).
16. W. M. Sinton, *Sky and Telescope,* **19,** 348 (1960).
17. R. J. Coates, Report of NRL Progress, p. 24. U.S. Naval Research Laboratory, Washington, D.C., July (1959).
18. N. A. Amenitskii, R. I. Noskova and A. E. Salomonovich, *Astr. J., Moscow* **37,** 185 (1960).
19. P. G. Mezger and H. Strassl, *Planet. Space Sci.* **3,** 213 (1959).
20. A. Giraud, Radiation Laboratory Report No-03544-I-T, Rensselaer Polytechnic Institute (1960).
21. F. H. Mitchell, R. N. Whitehurst and J. B. Dozier, *Bull. Amer. Phys. Soc.* Series 2 **2,** 282 (1957); *A Radio Study of the Sun and Moon at Millimeter Wavelengths.* Department of Physics, University of Alabama (1958).
22. A. E. Salomonovich, *Astr. J., Moscow* **35,** 129 (1958). (English translation: *Soviet Astron.* AJ, **2,** 112 (1959)).
23. J. P. Hagen, Report 3504 U.S. Naval Research Laboratory, Washington, D.C. (1949).
24. M. R. Zelinskaya, V. S. Troitskii and L. I. Fedoseev, *Astr. J., Moscow* **36,** 643 (1959).
25. C. J. Grebenkemper, Report 5151. U.S. Naval Research Laboratory, Washington, D.C. (1958).
26. V. S. Troitskii and M. R. Zelinskaya, Astr. J., Moscow **32,** 550 (1955).
27. M. R. Zelinskaya and V. S. Troitskii, *Trans 5th Congr. on the Questions of Cosmogony* p. 99. Akad. Nauk S.S.S.R. (1956).
28. N. L. Kaidanovskii, M. T. Turusbekov and S. E. Khaikin, *Trans. 5th Congr. on the Questions of Cosmogony, March,* 1955, p. 347. Akad. Nauk S.S.S.R. (1956).
29. G. Westerhout, *Bull. Astr. Insts. Netherlds.* **14,** 215 (1958).
30. C. L. Seeger, G. Westerhout and R. C. Conway, *Astrophys. J.* **126,** 585 (1957).

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 314-320. Printed in Great Britain.

METHODS OF CALCULATING ATMOSPHERIC DENSITIES FROM SATELLITE ORBIT DATA

G. V. GROVES*
University College, London

(*Received* 22 *February* 1961)

Abstract—Formulae relating atmospheric density to the rate of change of period of a satellite have been obtained by various workers for two asymptotic cases: (i) eccentricity e small ($<0{\cdot}3$) and (ii) e large ($>0{\cdot}012$). These results are reviewed, and in the overlapping region, $0{\cdot}012<e<0{\cdot}3$, the formulae are shown to be in agreement to the various orders of magnitude considered.

1. INTRODUCTION

An effect of atmospheric drag on the orbit of a satellite is to change its shape from an elliptical to a circular form. The orbital elements affected are the semi-major axis a and the eccentricity e, and both decrease with time. The term $a(1-e)$, however, which is the perigee distance, decreases only very slowly, and over a sufficiently short interval of time may be considered to remain constant. Hence the rates of change of a and e are not independent to a first order, and in deriving the atmospheric density it is adequate to consider the change in only one of these quantities.

In practice, the period of revolution T of a satellite and changes in period ΔT are readily observed, and from these the corresponding changes in a may be derived from

$$\Delta T/T = \tfrac{3}{2}\Delta a/a \qquad (1)$$

If Δ denotes the change in one revolution, then $\Delta T/T$ is $\dot{T}$, the secular rate of change of T with time. For strict accuracy, in evaluating T consideration needs to be given to the effect of the non-central part of the earth's gravitational field: this correction becomes important for high perigee satellites with small $\dot{T}$[1].

The relationship between Δa and air density ρ is obtained by integration over one revolution of the instantaneous rate of change of a, which is[2]

$$a^{-1}da/df = -k\rho a(1+2e\cos f+e^2)^{3/2}/(1+e\cos f)^2, \qquad (2)$$

where f is the true anomaly (Fig. 1) and k is the drag parameter SC_D/m. The corresponding relation in terms of the eccentric anomaly E is

$$a^{-1}da/dE = -k\rho a(1+e\cos E)^{3/2}/(1-e\cos E)^{1/2} \qquad (3)$$

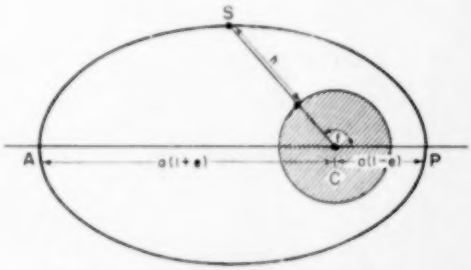


Fig. 1. Instantaneous satellite orbit.

For closed integration of (2) or (3), it is usually assumed that

(1) k, a, e remain constant during one revolution.

(2) ρ is a certain function of height h, say

$$\rho/\rho_0 = \exp[-(h-h_0)/H] \qquad (4)$$

where H is a constant, the atmospheric scale height.

(3) Either e is small compared with unity (case I), or that e is large compared with H/a (case II).

* Visiting scientist at the Air Force Cambridge Research Laboratories, Bedford, Mass.

Both cases have been investigated by various workers, and their results are here reviewed and compared. The notation used in the various original presentations has been preserved as far as possible, but some changes have been unavoidable.

2. CALCULATION OF AIR DENSITY WHEN $e<0{\cdot}3$ (CASE I)

When equation (2) is expanded in powers of e the integrals may be expressed in terms of Bessel functions of the first kind with imaginary argument,

$$I_n(x)=\frac{1}{\pi}\int_0^{\pi}\cos n\theta\exp(x\cos\theta)\,d\theta. \tag{5}$$

Use of this integral in calculating satellite drag effects was made by Roberson[3] with $x=ae(1-e^2)/H$, and by Leslie[4] and by Leeper[5] with $x=ae/H=c$, say. Sterne[6] followed this method and obtained the following formula, assuming an exponential density-height variation (equation (4)), but neglecting the oblateness and rotation of the atmosphere:

$$\Delta a=-2\pi k\rho_0 a^2\exp(-c)\,[b_0I_0(c)+b_1I_1(c)] \tag{6}$$

where

$$b_0=1+\frac{3}{2}e^2-\frac{e^3}{c}\quad\ldots \tag{7}$$

$$b_1=2e-\frac{3e^2}{2c}+e^3\left(1+\frac{2}{c^2}\right)\quad\ldots \tag{8}$$

ρ_0 is air density at perigee height. The formula given by Cook *et al.*[7] for the air density ρ_{00} at an initial perigee height is

$$\Delta a=-2\pi k\rho_{00}a^2\exp[(a_0-a-a_0e_0)/H]\{I_0+2eI_1+\tfrac{3}{4}e^2(I_0+I_2)+O(e^3)\} \tag{9}$$

This agrees with (6) to order e^2 in view of the identity

$$cI_2+2I_1=cI_0. \tag{10}$$

The effects of the oblateness and rotation of the atmosphere have been taken into account by Sterne[8] and Cook[9]. The rotation of the atmosphere results in a slight reduction in drag for satellites launched in an eastwards direction, i.e. having $i<90°$. If ω is the angular velocity of rotation of the atmosphere about the earth's axis its angular velocity about the pole of the orbital plane will be $\omega\cos i$. Hence the orbital component of satellite velocity relative to the atmosphere at a geocentric distance r is reduced by the ratio $r\omega\cos i/V$, where V is satellite velocity. The transverse component of velocity which is introduced will be very small compared with V, and hence the drag will be reduced by a factor of approximately[10]

$$F=\left(1-\frac{r\omega\cos i}{V}\right)^2 \tag{11}$$

For a nearly circular orbit r and V are approximately constant and F becomes, on neglecting terms of order e

$$F=(1-\tau)^2 \tag{12}$$

where

$$\tau=Td\Lambda\cos i \tag{13}$$

Td is the satellite's period of revolution expressed as a fraction of a sidereal day, and Λ is the ratio of the atmosphere's angular velocity to that of the earth. Taking $T_0=\frac{1}{15}$ $\Lambda=1$, we see that F ranges from about 0·85 to 1, according to the inclination. When k is replaced by

$$k'=k(1-\tau)^2, \tag{14}$$

the rotation of the atmosphere is taken into account with terms of order $e\tau$ neglected.

Taking atmospheric oblateness into account Cook obtains to order e

$$\Delta a=-2\pi k'a^2\rho_0\exp(-c-\tfrac{1}{2}\gamma\cos 2\omega)\,[I_0+2eI_1+\tfrac{1}{2}\gamma\cos 2\omega I_2] \tag{15}$$

where

$$\gamma=H^{-1}(\alpha-\beta)(q/\alpha)\sin^2 i, \tag{16}$$

α, β are equatorial and polar radii of the earth, $\alpha-\beta=21{\cdot}5$ km and i is the inclination of the orbital plane to the equator. q is perigee distance, $a(1-e)$. In deriving (15) it is assumed that $\frac{1}{2}\gamma$ is small, terms in $\frac{1}{4}\gamma^2$ being neglected. The oblateness effect is greater for orbits of high inclination. If i is near 90° $\frac{1}{2}\gamma$ may reach $\frac{1}{4}$ and terms in γ^2 could be at least as important as those in e^2.

Development of the theory to terms of order

e^3, γe^2 and $\gamma^2 e$ has been given by Sterne[8]. On account of the rotation of the atmosphere (3) becomes

$$\frac{da}{dE} = -k\rho a^2 \frac{(1+e\cos E)^{3/2}}{(1-e\cos E)^{1/2}} \left[1 - d\frac{1-e\cos E}{1+e\cos E} + O(d^2)\right] \quad (17)$$

where

$$d = \tau(1-e^2)^{1/2}. \quad (18)$$

On account of the oblateness of the atmosphere the density variation (4) becomes

$$\rho = \rho_0 \exp[-c(1-\cos E)] \left[1 + \frac{q_1 \sin^2 E}{(1-e\cos E)^2} + \frac{q_2 \sin^4 E}{(1-e\cos E)^4} + \theta\right], \quad (19)$$

where θ denotes odd terms and higher-order even terms, and

$$q_1 = (1-e^2)[-\gamma' \cos 2\omega + \tfrac{1}{2}\gamma'^2 \sin^2 2\omega] \quad (20)$$

$$q_2 = \tfrac{1}{2}(1-e^2)^2 \gamma'^2 \cos 4\omega \quad (21)$$

$$\gamma' = H^{-1}(\alpha=\beta)\sin^2 i. \quad (22)$$

With these generalizations (6) becomes

$$\Delta a = -2\pi k\rho_0 a^2 (1-d)^2 \exp(-c)[B_0 I_0(c) + B_1 I_1(c)] \quad (23)$$

where

$$B_i = b_i + b_i' d + O(d^2) \qquad (i=0,1) \quad (24)$$

$$b_0 = 1 + \tfrac{3}{2}e^2 - \frac{e^3}{c} + q_1\frac{e}{c}\left(4 - \frac{51}{2}\frac{e}{c}\right) + \frac{3q_2}{c^2}\left(1 - 24\frac{e}{c}\right)$$

$$b_0' = 4e^2 - \frac{2e^3}{c} + 4q_1\frac{e}{c}\left(1 - 9\frac{e}{c}\right) - 48q_2\frac{e}{c^3} \quad (25)$$

$$b_1 = 2e - \tfrac{3}{2}\frac{e^2}{c} + e^3\left(1 + \frac{2}{c^2}\right) + \frac{q_1}{c}\left[1 - \frac{8e}{c} + \frac{17}{2}e^2\left(1 + \frac{6}{c^2}\right)\right] - \frac{6q_2}{c^2}\left[\frac{1}{c} - 3e\left(1 + \frac{8}{c^2}\right)\right]$$

$$b_1' = 4e - \frac{4e^2}{c} + 2e^3\left(1 + \frac{2}{c^2}\right) + \frac{q_1}{c}\left[-\frac{8e}{c} + 12e^2\left(1 + \frac{6}{c^2}\right)\right] + 12q_2\frac{e}{c^2}\left(1 + \frac{8}{c^2}\right) \quad (26)$$

Neglected terms are of the order e^4, $q_1 e^3$ and $q_2 e^2$. Equation (15) is found to agree with (23) to the same order, when γ' is replaced $\gamma = \gamma'(q/\alpha)$. The factor q/α arises from the difference between the ellipticity of an equipotential surface through perigee and the ellipticity of the earth. For strict accuracy this factor should be included. In practice, however, q/α often differs from unity by an amount which is less than the accuracy with which γ' is known.

From (1) and (23), we have for perigee air density

$$\rho_0 = -\frac{\dot{T}}{3\pi ak(1-d)^2} \cdot \frac{\exp c}{B_0 I_0(c) + B_1 I_1(c)} \quad (27)$$

The air density at a height z above perigee, i.e. at $h = h_0 + z$ is then

$$\rho_a = \rho_0 \exp(-z/H) = -\frac{\dot{T}}{3\pi ak(1-d)^2} \cdot \frac{\exp[c(1-z/ae)]}{B_0 I_0(c) + B_1 I_1(c)} \quad (28)$$

In practice the value of H may not be well established, but the effect of its uncertainty on the accuracy of ρ_a may be minimized by choosing z so that $d\rho_a/dc = 0$. This condition requires, on neglecting terms of order e and γ in (28), $1 - z/ae - I_0'(c)/I_0(c)$; $I_0'(c) = dI_0/dc$, i.e. since $I_0'(c) = I_1(c)$,

$$z = ae[1 - I_1(c)/I_0(c)] \quad (29)$$

Hence the density at $h_a = h_0 + ae[1 - I_1(c)/I_0(c)]$ is

$$\rho_a = -\frac{\dot{T}}{3\pi ak(1-d)^2} \cdot \frac{\exp[cI_1(c)/I_0(c)]}{B_0 I_0(c) + B_1 I_1(c)} \quad (30)$$

3. CALCULATION OF AIR DENSITY WHEN $e > 0{\cdot}012$ (CASE II)

If the orbit is sufficiently elliptic, the major contribution to the orbital deceleration arises

in the region of perigee. In this case (1) or (2) may be expanded in terms of integrals of $\exp(-t^2)\,t^n$ for which asymptotic expansions may be introduced.

The required condition is that perigee should be lower than apogee by several scale heights, i.e.

$$ae \gg H \tag{31}$$

This method is usually applicable when $ae > 2H$, and hence $e > 0{\cdot}012$.

For an exponential density–height variation (equation (4)), it is found that

$$\Delta a = -k\rho_0 a^2 \left[\frac{2\pi H\,(1+e)^3}{ae\,(1-e)}\right]^{1/2} f \tag{32}$$

where f is unity when terms of order ae/H are neglected, and when atmospheric rotation and oblateness are also neglected. f may be expressed in terms of

$$\alpha_2 = ae\,(1-e)/2H\,(1+e) \tag{33}$$

as

$$f = 1 + \varepsilon'/\alpha_2 + O(\alpha_2^{-2}) \tag{34}$$

ε' has been evaluated[11] taking into account atmospheric oblateness and variation in scale height with altitude by replacing (4) by

$$\ln \rho/\rho_{00} = -(h-h_{00})/H_1 + (h-h_{00})^2/H_2^2 \tag{35}$$

When this is done (34) holds with (33) replaced by

$$\alpha_2 = \beta_1\gamma_2 - \beta_2\gamma_1^2 \tag{36}$$

where

$$\beta_1 = H_1/H \quad ; \quad 1/H = 1/H_1 - 2\,(h_0 - h_{00})/H_2^2 \tag{37}$$

$$\beta_2 = H_1^2/H_2^2 \tag{38}$$

$$\gamma_1 = \gamma_0 \sin 2\omega \quad ; \quad \gamma_0 = H_1^{-1}\,(\alpha-\beta)\sin^2 i \tag{39}$$

$$\gamma_2 = \gamma_1' + ae\,(1-e)/2H\,(1+e) \quad ; \quad \gamma_1' = \gamma_0 \cos 2\omega \tag{40}$$

h_{00} is an arbitrary height and h_0 is perigee height.

In order to exhibit the additional terms in (34) more clearly the expansion may be developed in terms of

$$\nu = ae\,(1-e)/2H\,(1+e) \tag{41}$$

It is then found that

$$f = 1 + (\tfrac{3}{4}H^2/H_2^2) + (\delta/\nu) + O(\nu^{-2}, H^4/H_2^4) \tag{42}$$

where

$$\delta = [\tfrac{1}{16}\,(1-8e+3e^2)/(1+e)^2] - \gamma_1^2\,(\beta_2 - \tfrac{1}{4}\beta_1^2) - \tfrac{1}{2}\beta_1\gamma_1'\,(1+\tfrac{3}{4}\beta_2\beta_1^{-2}) + O(\beta_2^2) \tag{43}$$

Equation (42) shows that the contribution to f from the variation of scale height with altitude is

$$R = 1 + \tfrac{3}{4}H^2/H_2^2 + O(H^4/H_2^4) = 1 + \tfrac{3}{8}dH/dh_0 + O(dH/dh_0)^2 \tag{44}$$

by (37).

The factor R has been investigated numerically by Jacchia[12], who finds that (i) R is practically independent of H, and (ii) $R-1$ is nearly proportional to the scale-height gradient for any given value of e. These conclusions are in accord with formula (44). Table 1 reproduces the values of Jacchia and compares them with equation (44).

Table 1. Values of R for $H = 63{\cdot}78$ km (Jacchia[12] and from equation (44))

e	dH/dh_0		e	dH/dh_0	
	0·1	0·2		0·1	0·2
0·00	1·000	1·000	0·20	1·039	1·083
0·01	1·032	1·061	0·40	1·039	1·082
0·02	1·050	1·100	0·60	1·039	1·082
0·05	1·044	1·095	1·00	1·039	1·082
0·10	1·040	1·086	equation (44)	1·0375	1·075

The effect of atmospheric oblateness is taken into account by the γ_1^2, γ_1' terms in (43). These increase $|\Delta a|$ by a fraction

$$\lambda = (\gamma_1^2\beta_1^2 - 2\gamma_1'\beta_1)/4\nu \tag{45}$$

on neglecting the $\gamma_1\beta_2$ and $\gamma_1'\beta_2$ terms, which arise from the cross-effects between atmospheric oblateness and non-constancy of scale height and give a somewhat smaller contribution. By (37) to (40),

$$\gamma = \tfrac{1}{2}\,(\nu H)^{-1}\,(\alpha-\beta) \sin^2 i\,[\tfrac{1}{2}H^{-1}\,(\alpha-\beta)\sin^2 i \sin^2 2\omega - \cos 2\omega] \tag{46}$$

The term $\frac{1}{2}H^{-1}(\alpha-\beta)\sin^2 i$ is usually small compared with unity as mentioned above. Hence the second term in λ predominates and leads to a reduction in $|\Delta a|$ of $\frac{1}{2}(\nu H)^{-1}(\alpha-\beta)\sin^2 i\cos 2\omega$. This reduction is greatest at $\omega=0, \pi$, i.e. when perigee is over the equator. The curvature of the atmosphere at perigee is then greatest, leading to a smaller drag effect. When perigee is nearest to the poles, the curvature of the atmosphere is least and the change in a revolution is greatest. The effect of the oblateness of the atmosphere increases as the eccentricity decreases, since ν then decreases; and by (41) we have approximately

$$\lambda=-(ae)^{-1}(\alpha-\beta)\sin^2 i\cos 2\omega \qquad (47)$$

for $ae\gg H$. λ may therefore lie within the range of ± 20 per cent.

It should be remarked that this variation in the drag effect with position of perigee arises solely from the curvature of the atmosphere. Another and somewhat larger variation arises from the change in perigee height on account of the oblateness of the earth, so that when perigee is over the equator, perigee height is correspondingly less and the drag is greater. This latter variation is taken into account by the variation of ρ_0 in the above equations.

The effect of atmospheric rotation, as well as that of oblateness, is included in the analysis of Sterne[8], but the scale height is assumed constant. It is found from (17) and (19) that

$$f=\left(1-d\frac{1-e}{1+e}\right)^3\left[1+\frac{f_1}{8c}+\frac{9f_2}{128c^2}+\ldots.\right] \qquad (48)$$

where

$$f_1=[(1-8e+3e^2)/(1-e^2)]-[16ed/(1+e)^2]+[8q_1/(1-e)^2] \qquad (49)$$

$$f_2=\frac{3-16e+50e^2+16e^3-5e^4}{3(1-e^2)^2}-\frac{16}{3}q_1(1+18e)+\frac{128}{3}q_2(1+4e)+\left[\frac{32e(e^2+4e-1)}{3(1+e)^2(1-e)^2}-\frac{256q_1e}{3(1+e)^2}\right]d \qquad (50)$$

Terms in q_1e^3 and q_2e^2 have been neglected in f_1, and terms in q_1e^2 and q_2e^2 in f_2. It is found that (42) and (48) agree, within their common orders of accuracy. (42) includes terms depending on the variation of scale height, while (48) includes the effect of the rotation of the atmosphere and terms of order $1/c^2$.

By (1) and (32)

$$\rho_0=-\frac{\dot{T}}{3k}\left[\frac{2e(1-e)}{a\pi H(1+e)^3}\right]^{1/2}\Big/ f \qquad (51)$$

The air density at height $h_a=h_0+z$ is then

$$\rho_a=\rho_0\exp[(-z/H)+(z^2/H_2^2)]=-\frac{\dot{T}}{3k}\left[\frac{2e(1-e)}{a\pi H(1+e)^3}\right]^{1/2}\exp(-z/H) \qquad (52)$$

on neglecting first-order quantities. To minimize the effect of errors in H, we choose z, so that $d\rho_a/dH=0$. This gives $z=\frac{1}{2}H$ (King-Hele[13], Whitney[14]). Hence (52) becomes, by (51),

$$\rho_a=-\frac{\mu\dot{T}}{k}\left[\frac{e(1-e)}{aH(1+e)^3}\right]^{1/2}(1+\tfrac{1}{4}H^2/H_2^2)/f \qquad (53)$$

where

$$\mu=\tfrac{1}{3}(2/\pi)^{1/2}\exp(-\tfrac{1}{2})=0{\cdot}1613. \qquad (54)$$

4. COMPARISON OF CASES I AND II IN OVERLAPPING REGION $0{\cdot}012<e<0{\cdot}3$

When the condition $ae\gg H$ is applied to case I, c is large and the asymptotic forms for $I_n(c)$ may be introduced

$$I_0(c)=(2\pi c)^{-1/2}\exp c\left(1+\frac{1}{8}c^{-1}+\frac{9}{128}c^{-2}+\ldots\right)$$
$$I_1(c)=(2\pi c)^{-1/2}\exp c\left(1-\frac{3}{8}c^{-1}-\frac{15}{128}c^{-2}+\ldots\right)$$
$$I_2(c)=(2\pi c)^{-1/2}\exp c\left(1-\frac{15}{8}c^{-1}+\frac{105}{128}c^{-2}+\ldots\right) \qquad (55)$$

From (6), the following asymptotic formula is obtained for $\dot{T}$ (Sterne[15, 16])

$$\dot{T}=\tfrac{3}{2}\Delta a/a=-3k\rho_0 a(\pi/2c)^{1/2}f(c,e), \qquad (56)$$

where

$$f(c,e)=1+2e+\tfrac{3}{2}e^2+\left[\left(1-6e-\frac{21}{2}e^2\right)/8c\right]+$$

$$+\left[\left(9-30e+\frac{171}{2}e^2\right)/128\,c^2\right]+O(e^3,c^{-3}) \quad (57)$$

From equation (9), the corresponding formula for ρ_0 is obtained, on putting $a_0=a$, $\rho_{00}=\rho_0$ (Cook *et al.*[7])

$$\rho_0=-\frac{T}{3k}\left(\frac{2e}{\pi aH}\right)^{1/2}\left[1-2e+\frac{5}{2}e^2-\frac{H}{8ae}\left(1-10e+\frac{7H}{16ae}\right)+\right.$$

$$\left.+O\left(e^3,\frac{H^3}{8a^3e^3}\right)\right] \quad (58)$$

From (30) and (4), the density at $h_a=h_0+\frac{1}{2}H$ is

$$\rho_a=-\frac{\dot{T}}{3\pi ak\,(1-d)^2}\cdot\frac{\exp(c-\frac{1}{2})}{B_0I_0(c)+B_1I_1(c)} \quad (59)$$

For large c this becomes by (55)

$$\rho_a=-\frac{\mu\dot{T}}{k}\left(\frac{e}{aH}\right)^{1/2}/g, \quad (60)$$

where

$$g=(1-d)^2\left[B_0\left(1+\frac{1}{8}c^{-1}+\frac{9}{128}c^{-2}+\ldots\right)+\right.$$

$$\left.+B_1\left(1-\frac{3}{8}c^{-1}-\frac{15}{128}c^{-2}\ldots\right)\right] \quad (61)$$

i.e. by (25), (26)

$$g=(1-d)^2\,(g_0+g_1d) \quad (62)$$

$$g_i=h_{i0}+(h_{i1}/8c)+h_{i2}/128\,c^2 \qquad (i=1,2) \quad (63)$$

$$h_{00}=1+2e+\tfrac{3}{2}e^2+e^3 \quad (64)$$

$$h_{01}=\left(1-6e-\frac{21}{2}e^2-11e^3\right)+8q_1\left(1+4e+\frac{17}{2}e^2\right) \quad (65)$$

$$h_{02}=3\left(3-10e+\frac{57}{2}e^2+75e^3\right)-48q_1\left(1+20e-\frac{153}{2}e^2\right)+384\,q_2\,(1+6e) \quad (66)$$

$$h_{10}=2e\,(2+2e+e^2) \quad (67)$$

$$h_{11}=-2e\,[(6+14e+11e^2)-16q_1\,(1+3e)] \quad (68)$$

$$h_{12}=-2e\,[3\,(10-38e-75e^2)+32q_1\,(15+81e)-768q_2] \quad (69)$$

For case II (53) has the form of (60) with g replaced by

$$g'=\frac{(1+e)^{3/2}}{(1-e)^{1/2}}\left(1-\tfrac{1}{4}\frac{H^2}{H_2^{\,2}}\right)f \quad (70)$$

On introducing the condition that e is small, neglecting $H^2/H_2^{\,2}$ and taking f from equation (48), g' is found to reduce to g as given by (62) to (69) to orders e^3, q_1e and q_2e, these being the orders of accuracy of f_2 (equation 50).

Alternatively if $H^2/H_2^{\,2}$ is retained and f is taken from (42), g' reduces to

$$g'=\left(h_{00}+\frac{h_{01}}{8c}+\ldots\right)+\frac{H^2}{H_2^{\,2}}\left(\tfrac{1}{2}h_{00}+\frac{h_{01}'}{8c}+\ldots\right) \quad (71)$$

where by (20)

$$h_{01}'=\tfrac{1}{4}\left(1-6e-\frac{21}{2}e^2-11e^3\right)-\left(1+4e+\frac{15}{2}e^2\right)(4\gamma'\cos 2\omega+17\gamma'^2\sin^2 2\omega) \quad (72)$$

i.e. (71) agrees with (62) to the orders of accuracy considered. On combining the terms in $H^2/H_2^{\,2}$ in (71) with those in (62), we have

$$g=(1-d)^2\,[g_0+g_1d+h_1\,(H^2/H_2^{\,2})] \quad (73)$$

where

$$h_1=\tfrac{1}{2}h_{00}+h_{01}'/8c. \quad (74)$$

5. SUMMARY AND DISCUSSION

Various formulae for the calculation of air density from satellite orbits have been reviewed and compared and shown to be in agreement to the orders of accuracy considered.

The quantities, concerning which approximations of smallness have been introduced, are

(i) eccentricity e

(ii) the ratio of scale height to eccentricity, or H/ae

(iii) ratio of oblateness of earth to scale height, or $\frac{1}{2}(\alpha-\beta)\,H^{-1}\sin^2 i$

(iv) ratio of angular velocity of atmosphere to that of satellite, or d

(v) the gradient of scale height with altitude, or H^2/H_2^2.

The order of magnitude of H^2/H_2^2 may be taken to be 0·05, and d is roughly the same for low inclination orbits. Hence, since (iv) and (v) have been taken into account to the first order, air density calculations may in general be made to an accuracy of better than one per cent on account of these effects. To maintain one per cent accuracy (iii) would need to be taken into account to the second order, particularly with high inclination orbits, and formulae are available for this. Case I, which treats e small, will be accurate to one per cent for e less than 0·3 if terms of order e^4 are neglected. Case II, which treats e as large and has been developed neglecting $(H/2ae)^3$, is accurate to one per cent if e exceeds about 0·012. In the range $0{\cdot}012 < e < 0{\cdot}3$, equation (60) will then be in error by less than one per cent on account of expansions with respect to (i) and (ii). The eccentricities of almost all satellites launched to date lie within this range. The few that lie outside may be analysed under either case I (equation (30)) or case II (equation (53)).

Above about 300 km the atmospheric density has been found to vary not only with height, but also with longitude (or solar elevation). This effect increases with height and at 700 km a factor of ten may be present between densities on opposite sides of the earth. For high perigee satellites, particularly those with small eccentricity such as 1960 ι (Echo), the diurnal air density bulge has an increasingly important effect on air density calculation and clearly needs to be considered.

Acknowledgement—The assistance of Mr. M. J. Davies, who checked through the equations in this paper and prevented several errors from occurring, is very gratefully acknowledged.

REFERENCES

1. D. G. King-Hele and D. M. C. Walker, *Symp. Aeronomy, Copenhagen, July 1960.*
2. W. M. Smart, *Celestial Mechanics.* Longmans Green, London (1953).
3. R. E. Roberson, *Jet Propuls.* **28,** 90 (1958).
4. D. C. M. Leslie, *Nature Lond.* **181,** 403 (1958).
5. E. Leeper, Rand Corporation, P.–1496, 24th September (1958).
6. T. E. Sterne, Special Report No. 11, Smithsonian Astrophysical Observatory (1958).
7. G. E. Cook, D. G. King-Hele and D. M. C. Walker, *Proc. Roy. Soc. A* **257,** 224 (1960).
8. T. E. Sterne, *Amer. Rocket Soc. J.* **29,** 777 (1959).
9. G. E. Cook, Tech. Note G.W.550. Royal Aircraft Establishment (1960).
10. D. G. King-Hele, *Nature, Lond.* **184,** 1267 (1959).
11. G. V. Groves, *Proc. Roy. Soc.* A **252,** 16 (1959).
12. L. G. Jacchia, Special Report No. 46, Smithsonian Astrophysical Observatory (1960).
13. D. G. King-Hele, *Nature, Lond.* **183,** 1224 (1959).
14. C. A. Whitney, Special Report No. 21, Smithsonian Astrophysical Observatory (1959).
15. T. E. Sterne, *Science* **127,** 1245 (1958).
16. T. E. Sterne, *Phys. Fluids* **1,** 165 (1958).

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 321-325. Printed in Great Britain.

ACCRETION AND EROSION ON THE SURFACE OF THE MOON

B. WARNER
University of London Observatory, Mill Hill Park, London

(*Received* 17 *April* 1961)

Abstract—By considering the accumulation of dust within a lunar crater (formed by the effects of accretion and erosion), formulae are derived that give (a) the radius of the crater floor and (b) the depth of the crater, as functions of the age of the crater. Comparison of these formulae with the observed relationships provides a means of dating the craters in an absolute sense.

1. INTRODUCTION

The hypothesis that the lunar maria are composed of dust was proposed by T. Gold in 1955[1]. Gold pointed out that the Moon is constantly accreting material in the form of micrometeorites, and the action of these and short-wavelength solar radiations on the lunar surface may cause an appreciable amount of erosion over intervals of time comparable with the age of the Moon. In this paper Gold's hypothesis will be applied to a particular problem of the lunar surface and accordingly its practicability will to some extent be tested.

The essential feature of Gold's hypothesis is that the dust formed by accretion and erosion is composed of particles so small that thermal and electrostatic effects would endow the dust layer with properties similar to that of a liquid. Thus wherever dust is formed it will flow downhill to the lowest point. In this way Gold explains the existence of the lunar maria and many other properties of the lunar surface.

Let us consider what happens to accreted and eroded material collected within the walls of a crater. Here the dust is trapped and cannot flow away towards the maria. Instead it will eventually travel downwards on to the floor of the crater. Over suitable periods of time the floor-level of the crater may be raised by large amounts.

The accretion rate of micrometeorites is 10^{-7} cm^3/cm^2/year[1]. In 10^9 years there will therefore be about 10^6 m^3 of dust collected within a crater of diameter 10 km (assuming that such a crater has been in existence that long). Spread over the area of the whole crater the dust will be 1 m thick, but if all the dust falls to the centre of the crater it may produce an increase in floor level (decrease in depth) of as much as 100 m.

Gold's estimate of the erosion rate is 10^{-4} cm^3/cm^2/year. Thus the floors of all craters may be considerably raised by the addition of a dust layer.

We will accordingly develop an elementary treatment of the accumulation of dust within the confines of a crater and later test our conclusions against observation.

2. THEORY

For convenience we will assume that the *initial* profile of all lunar craters is essentially parabolic. The effect of departures from this form will be mentioned in Section 5.

In Fig. 1 the radius of the crater is r and the depth of dust after a time t has elapsed is h ($t=0$ when the crater was formed). The radius of the floor (assumed flat) formed by the dust is ρ.

Let V be the volume of dust collected in the bottom of the crater after the time t, then

$$V=2\pi h^2 a$$

where a is the distance from vertex to focus of the parabolic profile. Therefore

$$\frac{dV}{dt}=4\pi a h\frac{dh}{dt}. \qquad (1)$$

Let D be the original depth of the crater, then

$$r^2=4aD. \qquad (2)$$

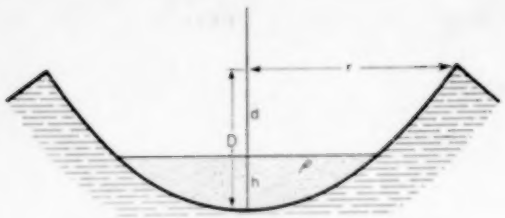


Fig. 1

If the accretion rate is u per unit area, per unit time,

$$\frac{dV}{dt} = u \,.\, \pi r^2. \tag{3}$$

From (1), (2) and (3):

$$h\frac{dh}{dt} = u \,.\, D.$$

Therefore if the present age of the crater is T and the present depth of dust on the crater floor is H:

$$\tfrac{1}{2}H^2 = u \,.\, D \,.\, T. \tag{4}$$

Now the radius of the floor ρ is given by $\rho^2 = 4aH$, whence from (2) and (4):

$$\rho^4 = 8uar^2T.$$

Putting $p = D/r$ we find

$$\rho^4 = \frac{2}{p}ur^3T, \tag{5}$$

where p is the initial depth/radius relation. We can find p from the observed depth/radius relation:

Let $q = \dfrac{d}{r}$ where d is the observed depth.

Then since

$$D = d + H,$$

$$p = q + \frac{H}{r}$$

and it follows from the geometry that

$$p = q + \frac{\rho^2}{r^2} \,.\, p.$$

Hence $$p = q\left[1 - \frac{\rho^2}{r^2}\right]^{-1} \tag{6}$$

We know q, ρ and r from observation.

Table 1 gives the resulting values of p for various crater diameter limits.

Table. 1.

Δ km	q	p
1–8	0·30	0·36
9–20	0·20	0·26
21–100	0·12	0·18

Putting

$$\delta = \text{diameter of floor} = 2\rho$$

$$\Delta = \text{diameter of crater} = 2r$$

we obtain from (5):

$$\begin{aligned} 1 < \Delta < 8 \quad & \delta^4 = 11u\Delta^3T \\ 9 < \Delta < 20 \quad & \delta^4 = 15u\Delta^3T \\ 21 < \Delta < 100 \quad & \delta^4 = 22u\Delta^3T \end{aligned} \tag{7}$$

Equations (7) give the expected floor diameters in terms of the measured quantity Δ, the parameter u, and the unknown age T.

Now it is well known that there is an apparent time sequence in the production of lunar craters. It is observed that no large craters are situated on top of smaller ones, so that the larger a crater is the older it is. There is some doubt[3] as to whether this relation is strictly correct, but for the present we assume that there is at least a statistical decrease in crater diameter with time during formation.

We may hope to find an empirical law such as

$$\Delta = kT^{\alpha}, \tag{8}$$

where k and α are positive constants.

Comparison of (8) with (7) shows that the slope of the observed $\log\delta - \log\Delta$ graph should be everywhere greater than 0·75. The actual slopes enables α to be determined.

We can determine the constants k and α by another, independent, method. So far we have used the relationship between depth and diameter only in finding p and for this purpose only approximate figures were needed. Now MacDonald[4] has provided the following empirical equation for lunar walled plains:

$$d = 0{\cdot}378\Delta^{\frac{1}{2}} - c \text{ km} \tag{9}$$

His values of c are a function of Δ, and with his published figures we find

$$d = 0{\cdot}378\,\Delta^{\frac{1}{2}} - 0{\cdot}089\,\Delta^{0{\cdot}89}\ \text{km}. \qquad (10)$$

We will find that equation (10) gives a second method of finding k and α.

Using (4), (2) and the fact that $H = D - d$, we find

$$d = \frac{r^2}{4a} - \left(\frac{uT}{2a}\right)^{\frac{1}{2}} . r. \qquad (11)$$

To reproduce the form of (10) we put

$$a = \lambda r^{3/2}, \qquad (12)$$

then

$$d = \frac{1}{4\lambda} . r^{\frac{1}{2}} - \left(\frac{uT}{2\lambda}\right)^{\frac{1}{2}} r^{\frac{1}{2}}.$$

Hence

$$\frac{1}{4\lambda} = 0{\cdot}378 \times 2^{\frac{1}{2}}$$

from the first term, i.e.

$$\lambda = 0{\cdot}47\ .\ \text{km}^{-\frac{1}{2}}$$

And

$$\left(\frac{uT}{2\lambda}\right)^{\frac{1}{2}} \left(\frac{\Delta}{2}\right)^{\frac{1}{2}} = 0{\cdot}089\,\Delta^{0{\cdot}59}$$

from the second term, i.e.

$$\Delta = 3{\cdot}7 \times 10^3 u^{1{\cdot}47} T^{1{\cdot}47} \qquad (13)$$

3. OBSERVATIONS

We have outlined in the previous section what is expected of lunar craters on the hypothesis proposed by Gold. Equation (13) gives the first result of combining theory with the observations by MacDonald, and we are left with a need for the observed δ, Δ relation.

Using the recently published *Photographic Lunar Atlas*[2] the writer has made measurements of δ and Δ. The values of Δ do not differ significantly from those given by Young[5]. For the large and medium sized craters the two measurements were both made under low lighting conditions. For the small craters the floor diameter was measured from the central grey spot apparent under high illumination, and the rim diameter measured under low illumination.

A plot of log δ against log Δ is given in Fig. 2. The scatter is partly due to errors in measurement and would probably be reduced by direct micrometric measurements on original plates.

It was expected, before constructing Fig. 2, that the slope of the graph would be larger than unity. This arises from an inspection of the photographs because for small craters δ/Δ is of the order of 0·5 whilst for many larger craters it is nearer 0·9.

4. COMPARISON OF THEORY AND OBSERVATIONS

A straight line solution is a good fit to the $\log\delta - \log\Delta$ graph. This yields the relationship

$$\delta = 10^{-0{\cdot}40}\Delta^{1{\cdot}12}. \qquad (14)$$

First let us consider the case of accretion alone. We put $a = 10^{-12}$ [km^3/km^2/year].

Comparison of (7) and (14) gives

$$\Delta \approx 10^{-6} T^{0{\cdot}7}\ \text{km for } \delta < 8\ \text{km}$$
$$\Delta \approx 10^{-7} T^{0{\cdot}7}\ \text{km for } \delta > 8\ \text{km}$$

where T is the age of a crater of diameter Δ km. Thus we find that a crater of diameter 10 km has to be at least 10^{10} years old. Similarly from (13) such a crater has an age of 10^{14} years. Since these ages are clearly impossible we conclude that accretion alone is insufficient to produce the observed $\log\delta - \log\Delta$ relation.

Erosion of the walls within a lunar crater will not work at a constant rate but will be proportional to the area of interior walls exposed to meteoric impact and solar radiation. Immediately after the crater is formed the whole of the crater interior will undergo erosion; when a layer of dust has been built up in the centre of the floor it will protect the rock below from further erosion and only the walls will remain exposed.

We will assume, as an approximation, that the average area exposed to eroding forces is equal to the area of the crater. This will be reasonable for small and medium-sized craters but will not approximate to the conditions for large craters.

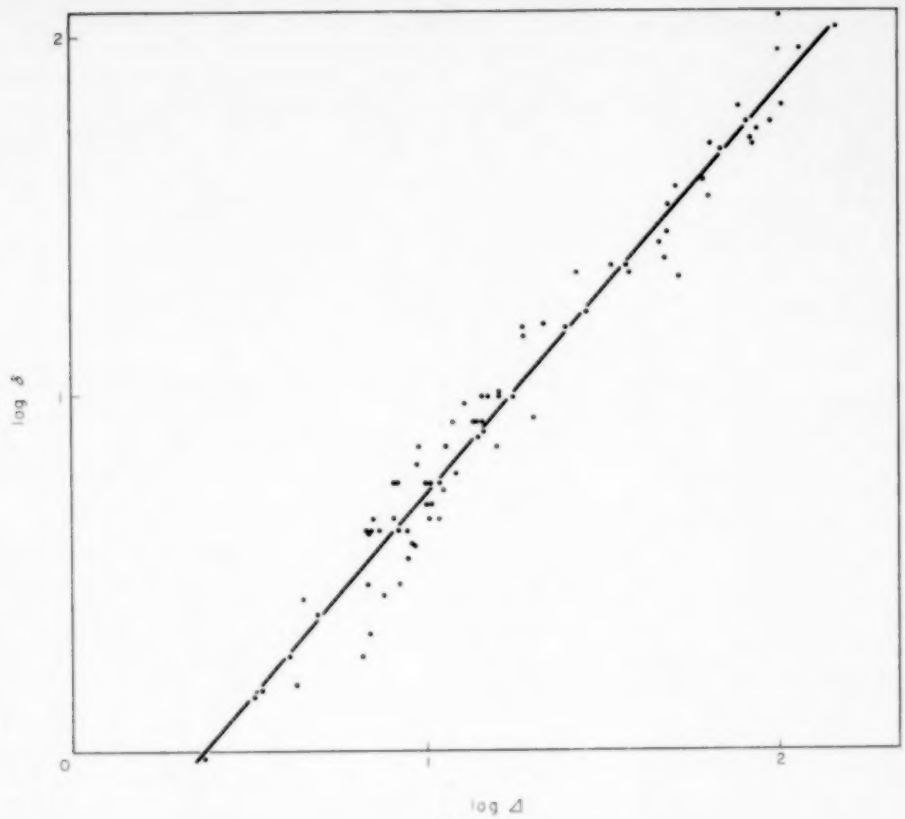


Fig. 2.

We can therefore use equations (7) and (13) for erosion, putting $u = 10^{-9}$ km^3/km^2/year.

Comparison of (7) and (14) yields

$$\Delta = kT^{0{\cdot}70},$$

where k is $10^{-4{\cdot}3}$, $10^{-4{\cdot}2}$, $10^{-4{\cdot}1}$ respectively for the increasing diameter groups. Combining these we obtain curve **a** in Fig. 3. This graph has been terminated at $\Delta = 30$ km because, as previously mentioned, the assumptions made do not apply to the larger craters. It can be seen that with the assumed value of u the smaller craters on the lunar surface have ages $\sim 10^7$ years.

The curve **b** in Fig. 3 is obtained by putting the erosion value of u into (13). The agreement between the two curves, constructed from independent observational measures, is strikingly good.

The slopes of the internal crater walls can be predicted from the theory of Section 2. We have seen that

$$\tan \alpha = \frac{\Delta}{4a}$$

where α is the slope at the top of the wall.

Using (12) and putting $\lambda = 0{\cdot}47$ km$^{-\frac{1}{2}}$ we find:

$$\tan \alpha = 1{\cdot}5\Delta^{-\frac{1}{2}},$$

whereas MacDonald[4] found:

$$\tan \alpha' = 2{\cdot}5\Delta^{-\frac{1}{2}},$$

where α' is the mean slope.

By putting $d=0$ and $T=\tau$ in (11) we find that age, τ, greater than which the craters will have been completely filled up with dust. Using (12) this gives:

$$\tau = 1{\cdot}9u^{-1}\Delta^4.$$

Thus we see that for accretion alone no craters will have been filled since the Moon was formed. For the erosion case there is the possibility that all craters $\gtrsim 4$ km diameter formed immediately after the Moon was formed will have been covered over.

5. CONCLUSION

The foregoing discussion has outlined a method by which lunar craters may be dated in an absolute sense. The ages derived are of a reasonable order of magnitude but are subject to the assumptions involved in the theory. Let us reconsider these assumptions.

We have used Gold's estimated rate of erosion, for which there is no experimental evidence. Changing this rate by any given factor changes the derived ages of the craters by the same factor. Thus a decrease of u by a factor of 30 from Gold's rate, or, which may be of more importance, an increase of u by a factor of 30 over the accepted accretion rate, both give ages of $\sim 3\times 10^9$ years for the larger craters.

To simplify the treatment we have assumed that the craters were initially paraboloidal. This probably breaks down for craters with diameters in excess of 30 km: these may have had flat floors of their own before any dust accumulated. Below 30 km the paraboloidal form is probably a fair approximation. If the craters were originally ellipsoidal their ages would be greater (and if hyperboloidal, less) than those given in Fig. 3.

The nature of the dark spots within small lunar craters, seen under high illumination, is not known. It has been proposed here that they are formed by collections of eroded and accreted material. Most craters with diameters greater than 30 km have obvious flat floors. Below this diameter the interior shadows are often not distorted as much as is to be expected by a flat floor. This may mean that the floors are slightly concave.

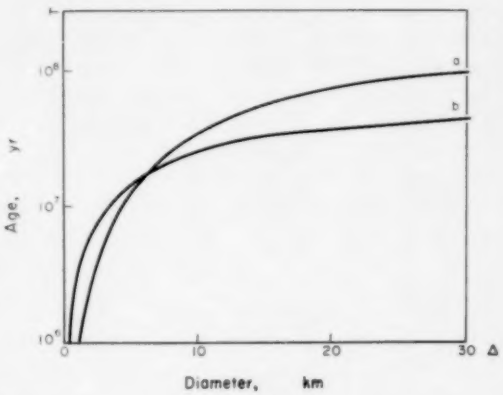


Fig. 3.

The scatter in Fig. 2 may give a measure of the dispersion in age for craters of a given diameter.

If the erosion rate were experimentally determined it would provide a test of the ideas given in this paper. Alternatively if the accretion rate was revised upwards by a factor of 30 or more then there would be no need to appeal to erosion. It may then prove possible to derive absolute ages for the lunar craters.

REFERENCES

1. T. GOLD, *Mon. Not. R. Astr. Soc.* **115**, 585 (1955).
2. G. P. KUIPER (ed.) *Photographic Lunar Atlas*, Univ. of Chicago Press, Chicago (1960).
3. B. WARNER, *J. Brit. Astr. Ass.* In press.
4. T. L. MACDONALD, *J. Brit. Astr. Ass.* **41**, 228 (1931).
5. J. YOUNG, *A Catalogue of Lunar Craters*, Arthur, (1953).

Planet. Space Sci. Pergamon Press 1961. Vol. 5, pp. 326-328. Printed in Great Britain.

RESEARCH NOTES

The dispersion of whistlers compared with the geomagnetic latitudes of their sources

(*Received* 27 *April* 1961)

At a station for whistler research outside Uppsala an investigation was carried out during three thunderstorm seasons (1958-60) with the special aim of discovering the relations between lightning discharges and the whistlers which followed them. The recordings were carried out during the summer seasons because we were better able then to check by direction finders both the regions of thunderstorm activity and the occurrence of individual lightning discharges.

During periods of whistler activity (called whistler situations) it was possible to make simultaneous recordings of the geographical location of the actual lightning discharge, the atmospheric wave-form emitted, the numbers of partial discharges and whistlers which followed.

The recordings obtained allow of treatment from different points of view. Some results have been presented in earlier issues of this Journal and elsewhere[1, 2, 3]. More detailed descriptions have been given in these earlier publications of the instrumental equipment, the recording procedures and the methods of analysis.

The special problem discussed in this article will be whether or not the geomagnetic latitude of the localities of the lightning discharges is of importance as regards the dispersion constants of the whistlers caused by them.

As has been mentioned, the geographical situations of the thunderstorms were obtained by direction finders. The results were checked on synoptical maps and showed a very good agreement. Cases which indicated a factor of uncertainty in the recordings were omitted.

The following determinations of the dispersion constants used are based on a considerable number of individual values (at least 10). For one determination of the dispersion constant in a situation the individual whistlers indicate small variations of ± 5 per cent from the mean value.

In Fig. 1 is reproduced a map with geomagnetic latitudes and longitudes. The calculated dispersion constants are placed on the map in the thunderstorm regions which produced the respective whistlers. The values within rectangles and circles represent a sample of different whistler situations in which several thunderstorms have produced whistlers during the same situation. The values that are not enclosed represent scattered cases obtained on different thunderstorm days and at different times around the clock.

From the map it follows that the most reliable results are obviously obtained in the same local situation and where the values are calculated from different thunderstorms. It is evident that a higher latitude of the lightning discharge region results in a higher dispersion constant. This is especially evident as regards the recordings marked within rectangles. The differences of latitude for the thunderstorm regions in these cases are 4° with the respective dispersion values of 140, 120 and 105. Also the other cases analysed show a marked decrease of the dispersion values from 185 at 60° geomagnetic latitude to about 100 at 42° latitude.

It follows from this that the geomagnetic latitude of the lightning discharges has a determining influence on the dispersion of the following whistlers. Therefore the locality of the receiving station has no influence on the dispersion. This is quite in accordance with the experimental results obtained by Morgan and Dinger[4]. They showed that two stations with a latitude difference of 5° recorded the same dispersion for the same whistlers. Recent investigations by Allcock[5] have shown that two stations with a north-southern distance of 2800 km between them can record the same whistlers with equivalent dispersions. In this connection it is of special interest to mention that we have observed how recordings on different days of lightning discharges developed in the same latitude result in different disper-

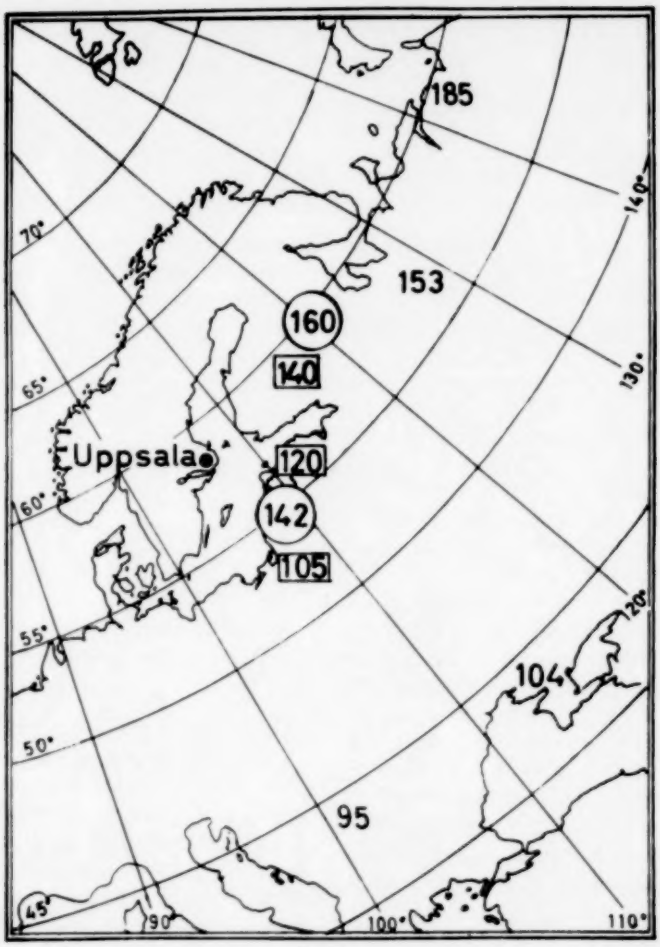


Fig. 1.

sions. Therefore some other factors must influence the dispersion constant. A marked seasonal variation has been shown by Helliwell[6]. Also variations around the clock have been observed.

The dispersion values given in Fig. 1 are low compared with the mean values calculated by Allcock[7]. This can largely be explained by the fact that our recordings were carried out during the summer seasons, when the dispersion values are low[6].

The main conclusion from our experimental analysis will be that the geomagnetic latitude of the lightning discharges is a determining factor with regard to the dispersion constant of the whistlers which follow the respective discharges Other factors might also have some influence.

To make these points clear it will be necessary to start a more extensive program, in which simultaneous recording by direction finders of the localities of lightning discharges and of the whistlers originating from them will be included.

Acknowledgement—The geophysics research reported in this note has been sponsored and supported by Contract AF 61(052)-07, Geophysics Research Directorate, Cambridge Research Center of the Air Research and Development Command, United States Air Force, through its European Office in Brussels.

REFERENCES

1. H. NORINDER and E. KNUDSEN, *Planet. Space Sci.* **1**, 173 (1959).
2. H. NORINDER and E. KNUDSEN, *Planet. Space Sci.* **5**, 46 (1960).
3. H. NORINDER and E. KNUDSEN, *Ark. Geofys.* **3**, 17, 347, Vetenskapsakademien, Stockholm 1961.
4. M. G. MORGAN and H. E. DINGER, *Nature, Lond.* **177**, 29 (1956).
5. G. Mc K. ALLCOCK, *Nature, Lond.* **188**, 732 (1960)
6. R. A. HELLIWELL, Communication at the XIII URSI Assembly in London 1960.
7. G. Mc. K. ALLCOCK, *J. Atmos. Terr. Phys.* **14**, 185 (1959).

HARALD NORINDER
EDGAR KNUDSEN

Institute of High Tension Research
Uppsala University

An estimate of the temperature in the exosphere

(*Received* 12 *June* 1961)

It was pointed out in a recent paper (Bates and Patterson[1], to be referred to as BP henceforth) that the steady-state vertical distribution of hydrogen atoms through the lower part of the thermosphere is a sensitive measure of the temperature $T(\infty)$ in the exosphere. Direct information on this distribution was not available when BP was written. Since then Purcell and Tousey[2] have published data on the equivalent width $W_{69°}(z)$ of the absorption core of the solar Lyman alpha line viewed at a zenith angle of 69° from altitudes z between 100 km and 200 km over White Sands at 7 hours local time on July 21, 1959. In particular, they report that

$$W_{69°}(200\text{ km}) = 0{\cdot}014\text{ Å} \qquad (1)$$

and

$$W_{69°}(100\text{ km}) = 0{\cdot}028\text{ Å} \qquad (2)$$

(which is beyond the linear part of the curve of growth). We shall use results of calculations described in BP to estimate $T(\infty)$ from these measurements.

Let $N_\zeta(z \mid H)$ be the number of hydrogen atoms per unit area above altitude z along a line at zenith angle ζ and let

$$r(z_1/z_2) = N_0(z_1 \mid H) \Big/ N(z_2 \mid H). \qquad (3)$$

According to BP r (100 km/200 km) is 2·6, 4·1, 8·5 and 16·4 when $T(\infty)$ is 1000°K, 1250°K, 1500°K and 2000°K respectively.

From measurement (1) and Fig. 2 (BP) it may be seen that $N_{69°}(200\text{ km} \mid H)$ is $3{\cdot}0 \times 10^{12}/\text{cm}^2$ and thence from equation (21) (BP) and Fig. 1 (BP) that

$$N_0(200\text{ km} \mid H) = 1{\cdot}8 \times 10^{12}/\text{cm}^2. \qquad (4)$$

With the aid of measurement (2) and equations (24) (BP) and (25) (BP), it is now possible to determine the equivalent width $W^L_{69°}$ which would arise if only the hydrogen atoms in the relatively dense and cool layer just above the 100 km level (cf. BP) contributed to the absorption core. A value of 0·022 Å is obtained. It follows from Fig. 3 (BP) that the slant number of hydrogen atoms in this layer, $N^L_{69°}$ is approximately $1{\cdot}1 \times 10^{13}/\text{cm}^2$. The corresponding vertical number N^L_0 is $3{\cdot}9 \times 10^{12}/\text{cm}^2$. Addition of the derived value of $N_0(200\text{ km} \mid H)$ yields

$$N_0(100\text{ km} \mid H) = 5{\cdot}7 \times 10^{12}/\text{cm}^2. \qquad (5)$$

Finally it is seen from (4) and (5) that

$$r(100\text{km}/200\text{ km}) = 3{\cdot}2. \qquad (6)$$

Comparison with the calculated values cited in the second paragraph suggests that on July 21, 1959, $T(\infty)$ was about 1100°K. This should be regarded as a mean in view of the time required for the distribution of hydrogen atoms to respond to thermal changes (cf. Table 6, BP). It is likely that $T(\infty)$ was less than 1100°K in the early morning (when the measurements were made) and was greater than 1100°K in the afternoon.

Acknowledgement—This research has been supported by the Air Force Cambridge Research Center, Geophysics Research Directorate, of the Air Research and Development Command, U.S. Air Force, through its European Office, under Contract No. AF61(052)–131.

REFERENCES

1. D. R. BATES and T. N. L. PATTERSON, *Planet. Space Sci.* (this issue, p. 257).
2. J. D. PURCELL and R. TOUSEY, Memoires in −8° de la Société Royale des Sciences de Liège **4**, 274 (1961).

D. R. BATES
T. N. L. PATTERSON

Department of Applied Mathematics,
The Queen's University of Belfast.

Planet. Space Sci. Pergamon Press 1961. Vol. 5. Printed in Great Britain.

BOOK REVIEWS

W. J. G. BEYNON (Ed.): **Some ionospheric results obtained during the International Geophysical Year.** Elsevier, Amsterdam; Van Nostrand, London, 1961. Pp. xi+399, 72*s*.

Although the IGY is sometimes referred to as a co-operative experiment on an unprecedented scale, it seems nearer the truth to think of it as a co-ordinated series of measurements of different kinds, designed to expose the underlying connections between many of the specialised fields of study in the rapidly expanding science of geophysics. Judged by the quantity and quality of the data which have already been collected in the World Data Centres, the IGY was a success, but the real value of the achievement will ultimately be judged by the uses to which the data have been put and the conclusions which have been drawn from them.

The reduction of the large volume of raw data into a form which can be more easily assimilated had been in progress for only nine months when the URSI/AGI Symposium was held in Brussels in September 1959. Nevertheless 60 short papers were presented and have now been published in order that a wider audience can be made aware of what has already been achieved. The papers deal mainly with concise descriptions of the results obtained in experiments having limited objectives, and represent some of the pieces of a mosaic rather than the complete picture of which they will ultimately form a part.

C. M. MINNIS

ROBERT JASTROW (Ed.): **The exploration of space.** Macmillan, New York, 1960. Pp. 160, 38*s* 6*d*.

This is a report of a symposium held in April 1959. The authors are distinguished experts, who are participating in the U.S. space program, and the papers cover the field of current space research apart from biology. The report was published in the *Journal of Geophysical Research* in 1959 and this re-publication is unlikely to be of great value to the specialist, because it has inevitably been overtaken by events in many places. But as an attractively illustrated (though there is a suggestion of bugs in some of the vehicles) collection of papers, ranging widely in their degree of technicality, this volume should be helpful to other scientists who wish to assess the uselessness and fascination of space research and the enthusiasm of its practitioners. A skim through this book might even clarify the non-scientist's impression of space research and he would enjoy some passages such as the list of criteria in order of importance from (1) reliability to (8) cost; this refers to instrumentation.

Some of the papers, on our own atmosphere, the trapped radiation and rocket astronomy, are concerned with actual achievements. Others describe our knowledge of planets and meteorites from ground observation. Two papers discuss the plasma and magnetic field in space itself, a subject in which there is just enough knowledge to launch exciting arguments, which are further boosted by disagreements on the basic physics.

J. W. DUNGEY

WILLIAM LILLER (Ed.): **Space astrophysics.** McGraw-Hill, New York, 1961. Pp. 268. Illustrated. $10.

This excellent book is a write-up of a series of lectures given during 1959-60 in the University of Michigan. The title is not fully descriptive. Not only does the book cover those aspects of astrophysics, solar physics and geophysics which are being studied from rockets and near-Earth satellites, but also instrumentation and stabilization are considered both from the engineer's and the physicists's point of view.

More than half is written by the men who have, for many years, been designing and making the payloads of the U.S. scientific vehicles. They speak with the authority of practical success, and the value of many of their comments to others who are embarking on this new type of work, cannot be over-estimated. Much is new or has only appeared either in a rather unusual periodical or in a conference report. Indeed, the list of references (more than 200 spread over 40 periodicals) is itself of considerable value. The discussions cover the successes and failures of the first generation of observational vehicles. They are extended to the design considerations of the current, more sophisticated payloads. Little is said of the more exotic next steps.

This type of book can never be complete; nevertheless one or two chapters might have been dropped, possibly to be replaced by one on the astronomical radio observations that can now be carried out from satellites.

It is acknowledged that the book has been hastily produced; there are several slips and the format is not perfect. Nevertheless the approach is correct in a rapidly developing subject—to get into print and circulate widely the results, ideas and experiences which are known only to a few busy men.

H. E. BUTLER